This report was funded by the Bonneville Power Administration (BPA), U.S. Department of Energy, as part of BPA's program to protect, mitigate, and enhance fish and wildlife affected by the development and operation of hydroelectric facilities on the Columbia River and its tributaries. The views in this report are the author's and do not necessarily represent the views of BPA.

For additional copies of this report, write to

Bonneville Power Administration
Public Information Center - CKPS-1
P.O. Bbx 3621
Portland, OR 97208

Please include title, author, and DOE/BP number from the back cover in the request

INTRODUCTION TO SURPH.1 ANALYSIS OF RELEASE-RECAPTURE DATA FOR SURVIVAL STUDIES

SURPH.1

Release Date: December 1994

Prepared by:

Center for Quantitative Science
School of Fisheries
University of Washington

Prepared for:

U.S. Department of Energy
Bonneville Power Administration
Environment, Fish and Wildlife
P. 0. Box 3621
Portland, OR 97208-362 1

Project Number **89-** 107
Contract Number **DE-BI79-90BP02341**

Contributors

This statistical software was developed as a team-effort involving biometricians, computer programmers, and biologists.

Authors

Steven G. Smith - Research Consultant

John R. Skalski - Principal Investigator

J. Warren Schlechte - Research Assistant

Annette Hoffmann - Research Assistant

Victor Cassen - Scientific Programmer

Acknowledgments

We wish to thank the following people for their support in developing this software:

- Pat Poe - Bonneville Power Administration

Dale Johnson - Bonneville Power Administration

Funding

Statistical theory and software development was funded by the Bonneville Power Administration under contract.

Contact Number: **DE-BI79-90BP02341**

Project Number: 91-051

Foreword

Program SURPH is the culmination of several years of research funded by Bonneville Power Administration to develop a comprehensive computer program to analyze survival studiesof fish and wildlife populations. Development of this software was motivated by the advent of the **PIT-**tag (Passive Integrated Transponder) technology that permits the detection of **salmonid** smolt as they pass through hydroelectric facilities on the Snake and Columbia Rivers in the Pacific Northwest. Repeated detections of individually tagged smolt and analysis of their **"capture-**histories" permits estimates of downriver survival probabilities. Eventual installation of detection facilities at adult fish ladders will also permit estimation of ocean survival and upstream survival of returning salmon using the statistical methods incorporated in SURPH. 1. However, the **utility** of SURPH.l far exceeds solely the analysis of **salmonid** tagging studies. Release-recapture and radiotelemetry studies from a wide range of terrestrial and aquatic species have been analyzed using SURPH.l to estimate discrete time survival probabilities and investigate survival relationships.

The interactive computing environment of SURPH.l was **specifically** developed to allow researchers to investigate the relationship between survival and capture processes and environmental, experimental and individual-based covariates. Program SURPH.l represents a significant advancement in the ability of ecologists to investigate the interplay between morphologic, genetic, environmental and anthropogenic factors on the survival of wild species. It is hoped that this better understanding of risk factors affecting survival will lead to greater appreciation of the intricacies of nature and to improvements in the management of our **wild** resources.

This BPA technical report **ia** an introduction to SURPH.l and provides a user guide for both the UNIX and MS-Windows@ applications of the SURPH software. The full **SURPH.1** documentation and diskette versions of the software are available from the authors or are accessible through the intemet at **ftp://opus.cqs.washington.edu/public/surph** or **http://www.cqs.washington.edu/~surph** .

Table of Contents

Table of Contents (cont.)

List of Figures

Page

List of Tables

Chapter 1

Introduction

There is increasing concern in the scientific community and in the public at large about natural and anthropogenic effects on the health of wild populations. Changes brought about by human activity can alter the ways animals interact with their environment. Consequently, it is important to understand the direct and indirect effects of the environmental milieu in which an individual animal finds itself.

A prime example is found on the Columbia and Snake Rivers in the northwestern United States. The desire to study the fate of salmon and steelhead populations native to these rivers was the motivation for the development of the SURPH methodology. Over the last century, major changes in river use have occurred on the Columbia and Snake Rivers. Sharp declines in salmon and steelhead runs have occurred concurrently with construction of hydroelectric projects, urbanization, and withdrawal of waters for irrigation. Hatchery production and mitigation projects—such as the construction of smolt bypass systems and adult fish ladders at the dams, smolt transportation programs and flow management-have been used in attempts to reverse the declines in salmon numbers and preserve salmon runs. However, very little is known about how environmental conditions, human actions, and mitigations may be affecting the survival of juveniles migrating to the ocean or adult salmonids returning to spawning streams.

Two technological developments have come together to help quantify survival relationships in the Columbia River Basin. The first new technology is the development of the PIT-tag (Passive Integrated Transponder) and the installation of PIT-tag detectors and slide-gate facilities at hydroprojects on the Columbia-Snake River systems. With this technology, juvenile salmon can be PIT-tagged, released and potentially detected at multiple dams as they migrate to the ocean. With proposed adult detector systems at various hydroprojects, these same fish might also be detected as they move upriver to spawn as adults. The result is the collection of capture history data that can be analyzed to estimate survival probabilities.

The second development is the advent of affordable high-speed computers that can numerically analyze increasingly sophisticated statist&l models of tag-release studies. This report describes statistical theory and computer software capable of not only extracting survival estimates but also characterizing relationships of survival rates with environmental **covariates** and the traits of individually marked fish and wildlife. The applications of the statistical models go far beyond the PIT-tag studies of salmon and steelhead on the Columbia and Snake Rivers. Applications of the methodology include the broader study of selection in wild populations based on the fates of individuals with differences in morphology, behavior, and genetics. Furthermore, using the statistical methods described in this report, field experiments can be conducted to test relationships between the survival of wild populationa and manipulative treatments applied over time or across the landscape. The methods presented in this report have been used to analyze tagging studies of not only salmon but also **molluscs,** insects, mammals, and avian species. The goal of this report is to convey the wide applicability and flexibility of analyses of tag-release studies to study survival relationships.

1.1 Background

Writing in 1652, **Izaak** Walton reported that Sir Francis Bacon studied salmon "by tying a **riband...in** the tail of some young **Salmons...and** then taking part of.them again, with the known mark, at the same place, at their return from the sea which is usually about six months later" (from Cormack 1970). This appears to be the first reference to the use of animal marking to study a population. It **would** be almost three hundred years before David Lack pioneered the estimation of survival rates from the recoveries of bands from marked birds that had died (Lack 1951). Since that time, much has been written on the use of bands, tags, or radiotags to estimate population parameters, such as survival rates and abundance, and to delimit movements (see Seber 1982 for review).

This report concentrates on methods that have been developed for estimation of survival probabilities and extensions of these models to investigate the effects of environmental factors and individual traits on survival. Animal marking studies can also be used to estimate other population parameters, such as abundance, density, and rates of movement between different habitats, but these methods are not covered here.

Statistical models for **survival** estimation have been developed for three types of tagging and resampling strategies. We will refer to all three strategies generically as tug-release sampling. **In** the first tag-release strategy, which we refer to as release-recovery, animals are captured, tagged with a number- or color-coded band, and released. The marked animals are subsequently resighted only once, usually upon death. Tagging may take place on one or several **occasions** throughout the duration of the study. The tag may identify the animals uniquely (e.g., the tag has a unique number) or may only indicate membership within a group of tagged animals (e.g., different groups are given tags of different colors). **In** Europe, ornithological studies usually investigate overall survival rates, and tags are collected from all possible mortality sources (Seber **1962, 1970,** and 1972; **Fordham** and Cormack 1970). **In** the United States, the majority of release-recovery models have been concerned with tag returns from waterfowl upon harvest by hunters (Brownie et al. 1985). Besides survival probabilities, release-recovery models must also include parameters for recovery or reporting rates. Program SURPH is ***not*** set up to analyze data from release-recovery studies.

In the second tag-release strategy, which we refer ***to as release-recapture, animals are tagged*** with a uniquely coded band or ring and released, just as for the release-recovery methods. But in release-recapture studies, the resampling occasions consist of the recapture of **tagged** animals while still alive. The group of animals caught on each resampling occasion may include both previously marked and unmarked animals. The recaptures of marked animals are recorded and the unmarked animals are usually marked, and then all of the animals are released back into the environment. Some models for release-recapture data, such as the "Jolly-Seber model" (Jolly 1965, Seber 1965) use the ratio of marked-to-unmarked animals in the resamples to estimate birth rates and total abundance, as well as survival rates and capture probabilities. Other models, such as those presented by Cormack **(1964), Burnham** et al. **(1987),** and Lebreton et al. **(1992),** use only the data from previously marked animals, and estimate only survival and capture probabilities. Release-recapture data can be analyzed using Program SURPH to estimate survival and capture rates based on an extension of the Cormack (1964) model.

The third strategy, ***termed known-fate, uses*** radiotags or other sampling techniques to track the fate of each marked individual. **In** terrestrial radiotelemetry investigations, animals are captured and fitted with small radio transmitters and released back onto the study site. Each transmitter has a unique signal. Resampling consists of triangulating on the radio signals until the animal is located and it is determined whether the animal is still alive. During the course of resampling, previously untagged animals may be captured and radiotagged to increase the sample numbers. Under ideal circumstances, each animal is relocated on each resampling occasion, obviating the need to include capture parameters in known-fate models (White 1983, White 1990). The establishment of upstream adult PIT-tag detectors with 100% scanning success at fish ladders holds the promise of providing known-fate data for returning adult salmon and steelhead on the Columbia-Snake River system. Other examples of known-fate studies include the study of sessile intertidal species. Program SURPH can be used to analyze known-fate data based on an extension of the White (1983) model.

For many years, the focus of tag-release models has been on the estimation of survival. The output of models ***we call survival-estimation*** models is a series of estimates of survival probabilities for the intervals between sampling events. Much of the early development of survival-estimation models for release-recover data was performed by Seber (1962, 1970, and 1972) and Cormack (1964). Fordham and Cormack (1970) used the models to estimate survival rates of Dominican gulls and included a statistical appendix by Cormack. Clobert et al. (1985) illustrated extensions of the Cormack model using data from studies of starlings, moths, and other animals. The monograph by Brownie et al. (1985) contains numerous models for analyzing tag returns from waterfowl hunters; this class of models has come to be known as Brownie models. A survival-estimation model for analyzing known-fate and radiotelemetry data was suggested by White (1983), who presented an example based on a tagging study of a population of elk. The models of Cormack (1964) and White (1983) form the basis for survival analyses in Program SURPH.

Survival-estimation models are seriously limited when it comes to assessing effects on survival or modeling survival relationships. The sequence of survival estimates from a tagging study are correlated, and there is no way to explicitly characterize relationships between survival probabilities and concomitant variables. Clobert et al. (1985) and Sandland and Kirkwood (1981) developed models with constraints on the survival and/or capture probabilities wherein survival or capture probabilities for two or more periods are equated and the common probability is estimated. However, researchers are not likely to be satisfied knowing only that per-period survival probabilities fluctuate or that the probabilities in different periods are equal. Rather, they will wish to relate the fluctuations and equalities to the **influence** of some environmental variables or traits of the individual animals. For example, Clobert et al. (1985) reported a study of a starling population ***(Sturnus vulgaris) in*** Belgium where the Cormack (1964) survival estimation model was used to analyze the release-recapture data. The estimates were then plotted against the average October-March temperature, and a clear linear relationship was suggested. Unfortunately, correlations among survival estimates induced by the estimation procedure itself invalidate the correlation coefficient or a regression line of the relationship. This shortcoming has led to the current interest in statistical models that explicitly represent the relationships between survival probabilities and environmental and experimental conditions or individual ***traits (covariates).*** We refer to such models ***as survival-effects*** models. **Survival-**effects models may also include relationships between capture probabilities and covariates.

From the early 1900's biologists have studied selection by attempting to establish a relationship between an explanatory trait and survival. In a study on praying mantises, **DiCesnola** (1904) found that color was important to survival when the insect inhabited an environment in which it was not camouflaged. The relationship was presented as percentages of green and brown preying mantises that survived in green and brown environments The results were extreme. By the end of the study, all the green preying mantises in a brown environment died and all the brown preying mantises in a green environment died. Thus, color of the insect was thought to be an agent of selection. However, no functional relationship was established.

Beginning in the **1980's,** biologists began using more sophisticated statistical analyses to find evidence for selection. In a study on zooplankton, Byron (1982) used Pearson's χ^2 goodness-of-fit test (Hogg and Tams 1983, pp. **414-420)** to analyze a contingency table of observed percent survival in several different **environments**. Thornhill (1983) applied linear regression to survival data on the scorpionfly where the response variable was an estimate of number of days survived and the explanatory variable was body size. **Schulze** and Folt (1990) compared Kaplan-Meier ***survival curves*** to treatment ***groups*** of a ***crustacean (Epischura*** *lacustris*). However, well into the 1980's, biologists relied mostly on descriptive relationships to provide evidence of selection (Endler 1986). The descriptive techniques ranged from plots of observed percent survival over tune for different groups of animals, to tables comparing percent survival in a group, to some characteristic of that group such as average size (Bantok and Bayley 1973).

Most of the published studies on selection have relied on either comparing traits averaged over a group of animals to survival in the group, or have relied on data where individuals could be tracked and their fates determined. Averaging traits over a group can average out the effect of selection or make the effects more difficult to detect (Clobert and Lebreton 1985). By tracking the fates of individuals, one can establish a relationship between survival and an individual trait such as body size, thereby studying selection at the individual level. Thus, there was a need for statistical methods that can be used to test hypotheses about selection at the individual level. For this reason, Program SURPH was developed to investigate both environmental and individual-based covariates.

Papers including survival-effects and capture-effects models for tag-release data are still infrequent in the literature, though the pace of new publication is increasing. The earliest was a paper by North and Morgan (1979) in which a relationship between the survival rates of first-year grey herons (*Ardea cinerea)* and the mean winter temperature was built into a model for release-recapture data. Poliock et al. (1984) extended a Jolly-Seber type model to relate capture probabilities to sampling effort in a Canadian lobster fishery. Pollock et al. (1984) also presented a general theory for relating capture probabilities to a wide variety of environmental and individual variables. In their monograph, Burnham et al. (1987) adapted a Cormack-type model to relate survival rates in desert tortoises to carapace length (pp. 361-371). The monograph also presented a general approach to assessing dose-response relationships (pp. 372-378). Clobert and his colleagues (Clobert et al. 1985, 1987, 1988; Lebreton et al. 1992) are currently extending the theory of survival effects models. A very general extension of the Cormack model has been developed that permits modeling of both survival or capture probabilities as functions of external variables. The model has been implemented successfully in a computer program called SURGE. Examples of the utility of the SURGE model are given in Clobert et al. (1988)

and Lebreton et al. (1992). Notable differences between Program SURPH and Program SURGE (Clobert et al. 1988) are the ability to analyze **known-fate** data; the types of link functions for modeling survival and capture relationships; incorporation of both group- and individual-based covariates; and the presence of formal methods of testing hypotheses and performing model diagnostics. Figures 1.1 and 1.2 summarize the range of survival-estimation and survival-effects models that have been devised for the three sampling strategies.

Program SURPH is capable of analyzing data on group- and individual-based covariates. Group-based covariates affect the overall survival of all individuals within a particular tag group, or population. Examples of group-based covariates include site- or **period-specific** environmental variables such as temperature or vegetation. Multiple-population studies are required to study site-specific variables. Individual-based covariates, such as body size, are measured on each individual and can be analyzed using a single population of tagged animals. Program SURPH can be used to model survival and capture as functions of both individual- and group-based covariates.

1.2 Capabilities of Program SURPH

Program SURPH is an interactive, mouse-driven statistical program developed for the **X-**Windows computing environment under the UNIX operating system for SUN@ workstations. SUN workstations were selected as the initial computing platform for SURPH because of their memory capacity and operating speed, important for iteratively solving the complex statistical models. Furthermore, SUN workstations have become a standard computing environment for numerous state and federal agencies dealing with salmon studies on the Columbia-Snake River systems. An analogous program SURPH-PC has also been developed for operation on IBM@-compatible personal computers using the Microsoft Windows* operating system. Both the UNIX and the Windows* versions of **SURPH** are available through the intemet (i.e., **ftp://opus.cqs.washington.edu/public/surph** or **http://www.cqs.washington.edu/~surph/surph.shtml**). Hard copies of the manual and diskettes of the program may be requested through the authors.

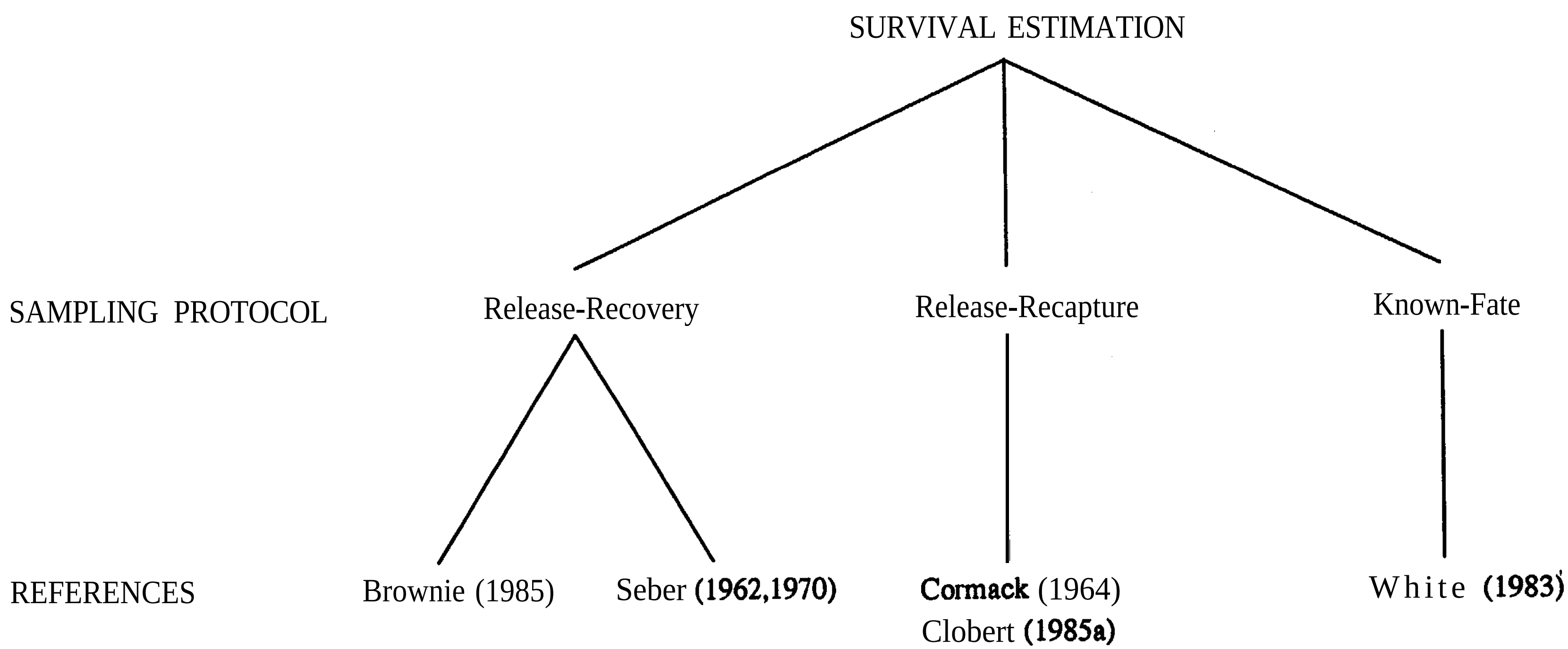


Figure 1.1 Diagram of available survival-estimation models for tag-release studies.

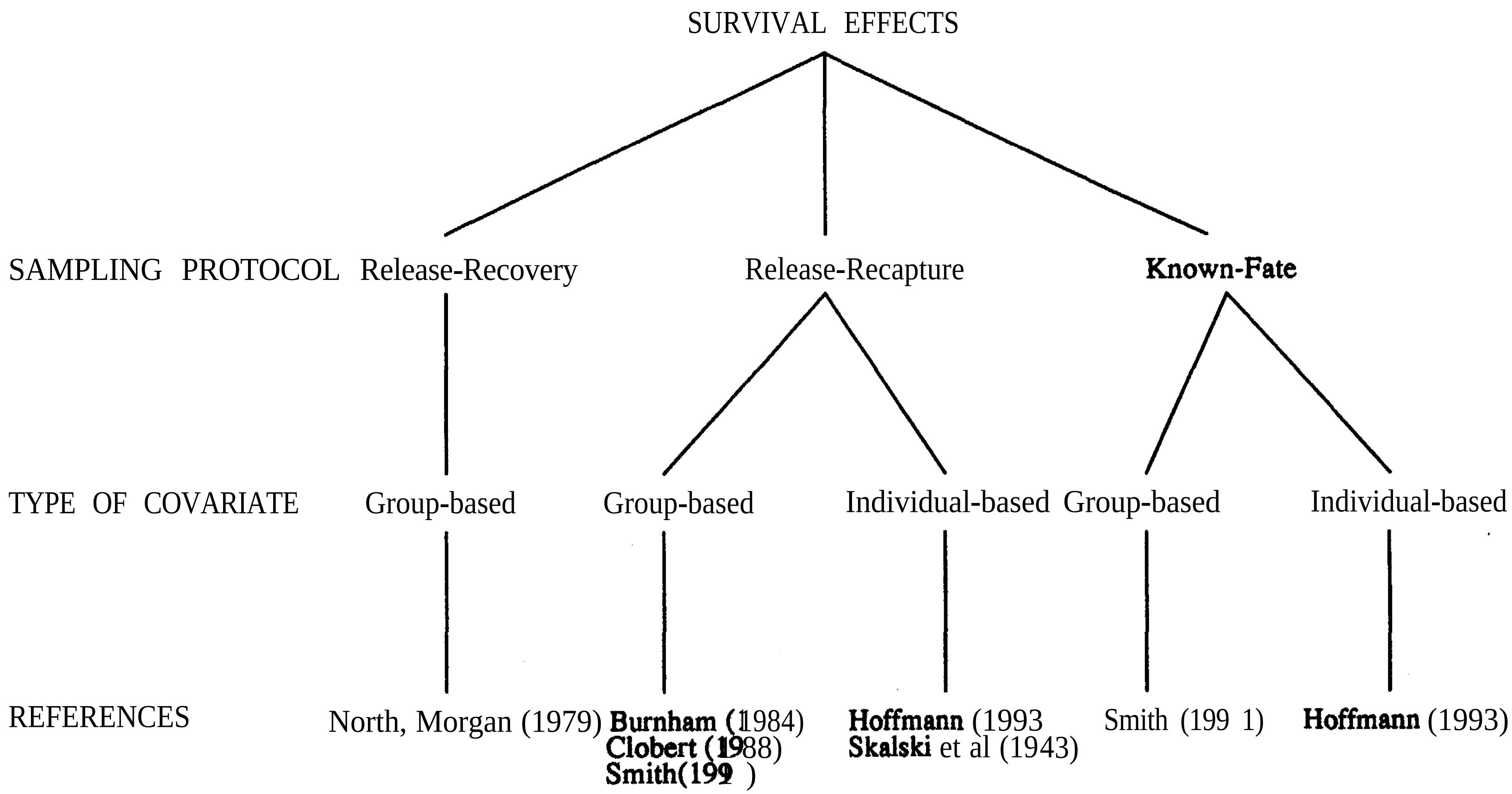


Figure 1.2 Diagram of available survival-effects models for tag-release studies.

Program SURPH has a wide range of capabilities to model and assess survival and capture relationships based on known-fate (Table 1.1) and release-recapture data (Table 1.2). The modeling and testing capabilities include a choice of link functions that relate the capture and survival probabilities to linear predictors (i.e., regression models). Capture and survival probabilities can be modeled using either a logistic-link function or a proportional hazards (log-log) link. In all cases, capture and survival probabilities can be characterized by period-, group-, and individual-specific effects.

Program SURPH, including both known-fate and release-recapture analyses, was devised to encompass the possible types of data sets that would ultimately be generated by Columbia-Snake River PIT-tag studies. Current juvenile PIT-tag detectors and slide-gate facilities generate release-recapture data that can be analyzed by Program SURPH. The eventual installation of adult PIT-tag facilities with 100% detection capabilities at upstream ladders will generate known-fate data that can also be analyzed by SURPH. Finally, with the prospect of both adult and juvenile PIT-tag facilities at Bonneville Dam (Figure 1.2), the juvenile tagging studies will generate data that is a composite of release-recapture and known-fate. The last interval from passage through Bonneville Dam as juveniles to redetection as returning adults provides known-fate data. For this reason, an unique capability of Program SURPH is the ability to model the last period as known-fate data to quantify and characterize ocean survival. 'Typically, release-recapture data will not permit estimates of survival in this last period. Rather, only the product of survival and capture probabilities can be estimated. Modeling in this case is usually limited to equality constraints among replicate release studies. The flexibility of Program SURPH will consequently allow analysis of a wide variety of terrestrial and aquatic tag-release studies.

Program SURPH is comprehensive in providing descriptive statistics, parameter estimates, model development, hypothesis testing, and model diagnostics (Tables 1.1 and 1.2). A useful feature of the interactive nature of Program SURPH is the ability to move easily between descriptive, estimation, testing, and diagnostic phases of tag data analysis to help identify the most descriptive yet parsimonious survival and capture models. This capability is reminiscent of some of the better statistical software for multiple regression. The parallels between SURPH

Table 1.1 Summary of features of Program SURPH for known-fate data.

FUNCTION	COMPUTATION
Descriptive Statistics	a. Kaplan-Meier (1958) nonparametric estimates of survival functions.
	b. Empirical c.d.f.'s of individual-based covariates
	c. Kolmogorov-Smirnov test of equality of covariate c.d.f.'s.
	d. Period-specific survival estimates and variances (White 1983).
Survival Analysis	a. Modeling individual- and group-baaed covariates.
	b. Parameter estimates and variances.
	c. Plots of relative risk.
	d. Plots of survival function versus individual- and group-based covariate values.
Diagnostics	a. Plots of parametric versus nonparametric survival estimates.
	b. Plots of residuals.
	c. Goodness-of-fit tests.

Table 1.2 Summary of features of Program SURPH for release-recapture data.

FUNCTION	COMPUTATION
Descriptive Statistics	a. M-arrays of capture data (Burnham et al. 1987).
	b. Cormack (1964) estimates of capture and survival probabilities and variances.
	c. Empirical c.d.f.'s of individual-based covariates.
	d. Kolmogorov-Smirnov tests of equality of covariate c.d.f.'s.
Modeling Capture Probabilities	a. Manly-Parr (1968) estimates of capture probabilities.
	b. Log-linear analysis of Manly-Parr (1968) estimates.
	c. Modeling individual- and group-based covariates.
	d. Likelihood ratio tests (LRT) of capture process.,
	e. Plots of capture process versus individual- and group-based covariate values.
	f. Parameter estimates and variances.
Modeling Survival Process	a. Modeling individual- and group-based wvariates.
	b. Tests of survival relationships using LRT and ANODEV.
	c. Parameter estimates and variances.
	d. Plots of survival process versus individual- and group-based wvariate values.
	e. Plots of relative risk.
Modeling Final Period	a. Modeling individual- and group-based covariates.
	b. LRT of parameter relationships.
Diagnostics	a. Plots of parametric versus nonparametric estimates of capture and survival probabilities.
	b. Plots of residuals.
	c. Goodness-of-fit tests.

software and regression software was intentional to rapidly acquaint investigators with this new methodology.

The software described in this report is intended for biometricians and biologists involved in the analysis of fish and wildlife survival studies. Investigators will be best prepared to use the methods described in this report if they have had a minimum of 2-3 college courses in applied statistics. These courses should include the general principles of statistical inference, regression analysis, and preferably experimental design and analysis of variance (ANOVA). In addition, readers should be familiar with tag-release theory. Valuable references on tag analysis may be found in Seber (1982), Clobert et al. (1985), Burnham et al. (1987), Lebreton et al. (1992), and Skalski and Robson (1992).

Chapter 2

Statistical Concepts and Methods

The theory underlying the assessment of effects on survival using SURPH is drawn from many difference sources; for example, tag-release theory, human epidemiology, maximum likelihood theory, and analysis of deviance. Elements of all these fields of study must be understood by any potential user of the SURPH software. This chapter is an introduction to the concepts on which the SURPH methodology is built.

Section 2.1 describes methods for modeling survival and capture processes for both continuous- and discrete-time data. Section 2.2 provides a general background on tag-release studies, including sampling protocols and concerns regarding the analysis of fish and wildlife survival studies. Section 2.3 covers the statistical method of maximum likelihood estimation. Section 2.4 discusses hypothesis testing using likelihood ratio tests and analysis of deviance. Section 2.5 discusses numerical procedures necessary to analyze complex survival models. Most biometricians should already be familiar with the material presented in Sections 2.3, 2.4, and 2.5 and may wish to familiarize themselves with details pertinent to wildlife tagging studies by reviewing the first two sections. Biologists, on the other hand, may wish to forego Sections 2.1 and 2.3 through 2.5 completely, focusing instead on the design and conduct of tagging studies discussed in Section 2.2. Investigators responsible for the analysis of tagging data will need to understand the information contained in Sections 2.1, 2.3, and 2.4.

2.1 Modeling Survival and Capture Relationships

The primary uses of Program SURPH are to evaluate and characterize relationships between measured covariates and survival probabilities. Modeling capture probabilities in release-recapture studies is usually a secondary objective. Section 2.1.1 explains the theory of continuous-time survival processes that underlies the discrete-time modeling of survival probabilities in SURPH. Section 2.1.2 discusses the range of mathematical functional forms that can be used to relate measured wvariates to survival and capture probabilities in tag-release models.

2.1.1 Continuous-time survival estimation

Let T be a nonnegative random variable representing the time elapsed from release to death of an individual from a homogeneous population. ***The survivor function*** for T, $S(t)$, *is* defined as the probability that T is equal to or greater than t. In other words, $S(t)$ is the probability an animal survives to at least time t:

$$S(t) = P(T \geq t), 0 < t < \infty. \tag{2.1}$$

If T is absolutely continuous, then $S(t) = 1 - F(t)$, where $F(t)$ is the cumulative density function for time to death. Note that $S(0) = 1$ and $S(\infty) = 0$.

The probability density function (p.d.f.) of T is

$$f(t) = \lim_{\Delta t \to 0^+} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = -\frac{dS(t)}{dt}. \tag{2.2}$$

By the definition of a p.d.f. for a nonnegative continuous random variable, $f(t) \geq 0$ for

$0 \leq t < \infty$, and $\int_0^\infty f(u)\,du = 1$. Also, $S(t) = \int_t^\infty f(u)\,du$. *The* function $f(t)$ describes the distribution of lifetimes in the population.

The instantaneous rate of failure at time $T = t$, conditional on survival to time t, is given by the ***hazard function*** and is defined as

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid (T \geq t))}{\Delta t} = \frac{f(t)}{S(t)} = \frac{-\frac{dS(t)}{dt}}{S(t)}. \tag{2.3}$$

From (2.3), $\lambda(t) = -\frac{d}{dt}\log S(t)$, because $\frac{d}{du}\log(x) = \frac{1}{x}\frac{dx}{du}$, so that by integrating we obtain

$$S(t) = exp\left(-\int_t^\infty \lambda(u)\,du\right). \tag{2.4}$$

and the p.d.f. of ***T can*** be written as

$$f(t) = \lambda(t)\,exp\left(-\int_t^\infty \lambda(u)\,du\right). \tag{2.5}$$

Rarely in the course of a survival study will a death time be observed for every single experimental subject. Some individuals may still be alive when the study is terminated; others may be "lost to follow-up" (i.e., they die before the end of the study, but the researchers are not aware of the time of death). When the death time of an experimental subject is not observed during the study, the death time is said to be ***censored.***

It is far more common to estimate the survivor function (Eq. (2.1)) than the hazard (Eq. (2.3)) or probability density function (Eq. (2.5)). However, the three functions are equivalent in that each uniquely specifies the probability distribution of ***T.*** Kalbfleisch and Prentice (1980, Chapter 3) give methods of parameter estimation when a parametric distribution is assumed for the death times. More commonly, however, a nonparametric method due to Kaplan and Meier (1958) is used to estimate the survivor function. The estimate produced by the nonparametric method is known ***as the Kaplan-Meier*** or ***product-limit*** estimate.

The Kaplan-Meier estimate is used to estimate the survivor function when the population is assumed to be homogeneous ***with*** regard to survival rates. The estimate is based on the ***empirical survivor function*** (***ESF***). If there are no censored observations in a sample of size ***n, the*** **ESF** is defined as

$$\hat{S}(t) = \frac{\text{Number of observations} \geq t}{n} \quad \text{for } t \geq 0 .$$

If all death times are distinct, the ESF is a step function that begins at $\hat{S}(0) = 1$ and decreases by l/n just after each observed death time. In general, if there are d *death times equal* to t, the ESF decreases by d/n just after t.

The Kaplan-Meier estimate can be modified to give an unbiased estimate of the survivor function when there are censored observations. Let $t_1 < t_2 < \ldots c\, t_k$ be the distinct observed death times in a study with sample size n $(n \geq k)$. "Ties" in death times are allowed; that is, more than one subject may die at each death time t_i, and we let d_i represent the number of deaths at time t_i $(i = 1, 2, \ldots. k)$. In addition to the death times, there are also censoring times C_i for individuals whose death times are not observed. Let c_i be the number of subjects censored at time C_i. The Kaplan-Meier estimate of the survivor function in the presence of censoring is defined as

$$S(t) = \prod_{i \in \{t_i, C_i\} < t} \frac{n_i - d_i}{n_i} . \tag{2.6}$$

where n_i is the number of individuals at risk at time t_i, i.e., the number of individuals alive and uncensored just prior to t_i ($n_i = n_{i-1} - d_{i-1} - c_{i-1}$). The Kaplan-Meier estimate reduces to the ESF if there is no censoring, because in that case, $n_, = n$ and $n_i = n_{i-1} - d_{i-1}$. Variance estimates for $\hat{S}(t)$ are given in Kalbfleisch and Prentice (1980) and Lawless (1982). Lawless (1982) also includes a section on estimation of the cumulative hazard function in his chapter on nonparametric estimation of the survivor function. For known-fate data, Program SURPH provides Kaplan-Meier estimates modified for periodic, or discrete-time, sampling and associated variance estimates (see Section 3.6.1).

Tag-release methods rarely provide the precise measurements of death times required for the

Kaplan-Meier method. However, some researchers have been able to pinpoint the days of death for moderately sized samples of tagged animals using radio transmitters and intense follow-up. Pollock et al. (1989a) report on one such study of a black duck population where fifty female ducks were fitted with radios. After tagging, the location and status (alive, dead, or missing) of each bird were recorded daily from the date of release (8 November 1983 to 14 December 1983) until 15 February 1984. The expense of locating a reasonable number of radiotagged animals on a daily basis is probably prohibitive for most researchers; it is far more common for radiotagged birds to be located on a less frequent basis; weekly, for example.

In a second paper published describing the adaptation of the Kaplan-Meier method to studies of wild populations, Pollock et al. (1989b) use an example where 52 bobwhite quail were radiotagged and then located weekly for five months. The Kaplan-Meier method may still be applied if the sampling is not done daily and exact death times are not known. To calculate the estimate, however, it is necessary to assign a death time to each animal that dies. The animals that die in the interval (t_i, t_{i+1}) could be assigned any death time between t_i and t_{i+1}, and the choice of assignment appears to be arbitrary. One solution is to assume that any deaths that occur during a sampling interval occur uniformly throughout the interval. Table 1 in Pollock et al. (1989b) implies the use of the death time $t_i + _1$ for all individuals dying in the interval $(t_i, t_i + _1)$ while their Figure 1 implies assignment of death times uniformly throughout the interval. Similar assumptions must also be made for the censoring times of individuals censored within the interval $(t_i, t_i + _1)$. When death times are discrete (i.e., not exact), different assumptions about the time of death or time of censoring may lead to different conclusions.

The Kaplan-Meier method, as originally devised, applies to a single group of subjects, all entering the study at the same time, which is called "time zero". In epidemiological studies, patients often do not enter the study at the same calendar time but all enter immediately after they have experienced a certain diagnosis or surgical procedure. Thus, time zero may refer to a different calendar time for each subject. It is assumed that the same survivor function applies to each subject following the event, regardless of the actual calendar time. In tag-release studies, there is no such "landmark" event in the life of a subject that "sets the clock." Rather, it is assumed that marked and unmarked animals experience the same hazards throughout the study period, so that if new animals are introduced to the study after the initial group is released, their time zero is assumed to be the same calendar time as the animals already released. This assumption is implicit in the adaptation of Kaplan-Meier methods to staggered-entry study designs proposed by Pollock et al. (1989a). The solution is to add the animals marked at time t_i to the "number at risk" at time t_{i+1}.

Release-recapture data cannot be analyzed properly using the Kaplan-Meier method. Unlike radio transmitters, passive tags do not provide death times of individual animals because dead animals are not subject to detection or recapture. Failure to recapture an animal may indicate either mortality or simply failure to recapture an animal that is alive. Release-recapture protocols rarely include searching for and recovering dead animals. The capture history data from release-recapture studies must be analyzed using likelihood models that include parameters for both capture and survival probabilities.

It is assumed under the Kaplan-Meier method that the population is homogeneous; that is, the same probability distribution for death times applies to all individuals in the biological population. However, this assumption is likely to be violated in any typical population. Because of inherent variation, some individuals are more likely to survive through any sampling interval than others, and the individual death times are influenced by intrinsic traits and exogenous factors. This is the incentive for using explanatory variables to model death times. Kalbfleisch and Prentice (1980) show how explanatory variables may be introduced to various parametric models, such as exponential, Gamma, or Weibull models. It is more common, however, not to assume a parametric model for survival times.

2.1.2 Discrete-time survival estimation

In studies of wild populations, it is often more interesting and appropriate to focus not on the entire survivor function, but on the probability of surviving from one sampling event to the next. The exact death times are no longer important; just the intervals within which deaths occurred. There are two reasons for this difference in focus from human epidemiology. First, death times often cannot be determined exactly in wild populations. Second, in manipulative studies such as pesticide trials, the treatment effect may influence survival for only a small period of time following the application of the treatment. It is -not necessary to study the entire survivor function to assess acute pesticide effects that quickly dissipate. Studying the entire function may even decrease power to detect effects occurring immediately after application.

There is a vast literature on tag-release models that includes **interval-specific** survival probabilities (survival-estimation models). The essential element common to all such models is the vector (S_1, S_2,..., S_K), where S_k (k = 1,..., K) is the conditional probability that an animal survives until sampling event k, given that it was alive just after event k-l. Periodic sampling naturally leads to consideration of conditional survival probabilities rather than instantaneous mortality rates, but there is still a continuous process at work. As before, let T be a random variable defined as the death time for an individual from a population with homogeneous survival rates. Let the hazard function be denoted $\lambda(t)$ and the survivor function $S(t)$. *The* conditional survival probability S_k can be written in terms of the continuous functions:

$$S_k = P(T > t_k | T \geq t_{k-1}) = \frac{S(t_k)}{S(t_{k-1})} = exp\left(- \int_{t_{k-1}}^{t_k} \lambda(u)\, du \right). \tag{2.7}$$

Now consider the simplest possible scenario for a single-population, single-release, release-recapture study, wherein the hazard function is constant for the duration of the study ($\lambda(t) = \lambda$ for all t) and the sampling events are evenly spaced (i.e., the sampling intervals are of equal length). Let the times or locations of the sampling events be denoted $t_1, t_2, t_3,..., t_K$ and let At be the width of the sampling intervals ($t_{k+1} - t_k$ = At for all k). In *this case, the survival* probabilities are equal:

$$S_k \equiv S = exp\left(- \int_{t_{k-1}}^{t_k} h du \right) = exp\,(-\lambda \Delta t) \quad , \tag{2.8}$$

where S is the survival probability common to all periods.

In any realistic situation, the hazard function is not likely to be constant throughout the study. Differences in the hazard function from interval to interval will be reflected in differences in the respective cumulative hazard functions

$$\left(\int_{t_{k-1}}^{t_k} \lambda\,(u)\, du \right)$$

for each interval, in turn resulting in differences in the survival probabilities. If the sampling intervals are of equal length, then differences in the survival probabilities can be attributed solely to interval-to-interval differences in the hazard function. However, survival probabilities will also vary across sampling intervals if the intervals are of unequal lengths, even if the hazard function is constant. For example, suppose a constant hazard function gives the probability of surviving one week as 0.7. Then the probability of surviving a two-week-long period is $0.7^2 = 0.49$. Thus, if the sampling events are not evenly spaced, differences in the survival probabilities due to a varying hazard function are completely confounded with differences due to uneven lengths of sampling intervals. It is impossible to separate the two causes.

2.1.3 Covariate effects on survival and capture processes

2.1.3.1 Types of covariates

The nature of the study design is also strongly dependent on the type of covariates investigated. SURPH is capable of analyzing two important classes of covariates. The first class is ***group-based.*** River flow, temperature, and habitat quality are examples of group-based covariates.

All the tagged individuals within a particular population experience effects of the same level of the group covariates. Group covariates have been incorporated into known-fate and release-recapture models by Clobert et al. (1987), Smith (1991), Lebreton et al. (1992), and Smith and Skalski (to appear). Group-based covariates may *be* further classified into ***time-invariant*** (fixed) and ***time-variant*** measurements.

If the covariates that characterize populations are fixed, replicate populations with differing covariate values are required to study the relationship between covariates and survival rates, just as different levels of the independent variable must be observed to perform a linear regression analysis. Time-variant, group-based covariates can be accommodated in a periodic-sampling model, if it is assumed that the value of the time-variant covariate is constant within a particular sampling period and changes only at the times of the sampling occasions. That is, there is a vector of covariates $\underset{\sim}{x}_k$ to which an individual is subjected from sampling occasion $k-1$ until occasion k.

Individual-based measurements are the second important class of covariates. Examples of individual-based covariates include measurements of the morphology of the animals, their genotypic or phenotypic characteristics, behavioral or territorial differences, gender, and handling history. Individual covariates like body size have been incorporated into known-fate models (Kalbfleisch and Prentice 1980, White and Garrott 1990, Pollock et al. 1989b) but have not previously been incorporated into release-recapture models.

Individual-based covariatcs can be either time-invariant (fixed) or time-variant. Time-invariant covariates have values that do not change (e.g., sex) or that change insignificantly during the course of the study. Fixed individual covariates are easier to model because the time of measurement does not matter; the relationship between the fixed covariate value and survival will be the same for all animals, regardless of when they entered the study.

In contrast, the relationship between time-variant covariates and survival depends on the time that the variable was measured. In a study with staggered entry of animals into the tagged population, covariate values are not comparable between animals measured at different times (i.e., upon entry into the study). For example, consider a pesticide study with weekly sampling occasions. A **first** set of tagged animals is assayed for blood contaminants just after the pesticide application that begins the study. The probability of surviving until the next sampling occasion can be modeled as a function of the contaminant level, because all animals were measured at the same time. Now suppose that at the second sampling occasion, a new set of animals is captured, measured for contaminants, tagged, and released. Unless all surviving animals in the first set are remeasured, the two sets cannot be used in the same regression model for survival in the next interval, because the covariate values are not comparable. The expected survival for an animal with a particular contaminant level measured on the second sampling occasion is not the same as that for an animal with the same level measured on the first occasion. When the covariates between groups of animals are not comparable, a separate analysis must be performed for each release group.

The relationship between survival and a time-variant individual-based covariate can be studied using a "single entry" design, wherein all tagged animals enter the study at the same tune. Often, a variable will be known to vary over time, but will prove to be very difficult to monitor; for example, the degree of smoltification of a juvenile **salmonid** as it migrates downriver. In many such cases, the solution is to measure the variable at the time of release and to treat it subsequently as time-invariant, being careful in the interpretation of effects in subsequent periods. When a time-variant covariate is treated as fixed, the observed relationship between the covariates and survival can be expected to degrade over time, as the initial measurement becomes less indicative of the **current** value of the covariate.

Some covariates can be treated as either group- or individual-based covariates. Examples include gender or age class, where the multiple individuals with the same covariate values can be defined as unique populations. In a single-site, tag-release study, estimation of regression coefficients will be identical regardless of how these covariates are treated in the data analysis. However, tests of signifkance will differ appreciably. If treated as group covariates, replicated sites, each with the population divided into subpopulations, would be required to show reproducible differences between classes of individuals. Treated as individual-based covariates, however, tests of significance are based on the inter-animal variation within a single population. This latter epidemiological approach to study design and analysis can be substantially more powerful in identifying significant survival effects. The epidemiological approach uses likelihood ratio tests **(LRT)** while group covariate effects must be tested using analysis of deviance (ANODEV) (see Section 2.4).

2.1.3.2 Quantities for comparing survival rates

This section describes several quantities that are used to characterize and contrast survival functions and probabilities. Implications to these quantities can help guide the selection of the appropriate link function.

The ***odds rario*** and ***the*** effect on ***relative risk*** (Hosmer and **Lemeshow** 1989) are both used to contrast survival probabilities of two individuals with covariate vectors $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$, respectively. If ***A*** is the event of interest and P(A) is the probability of event A, then the odds ratio for the two individuals ($OR(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$) is

$$OR(\underset{\sim}{x}_1 \underset{\sim}{x}_2) = \frac{P(A|\underset{\sim}{X} = \underset{\sim}{x}_1) / (1 - P(A|\underset{\sim}{X} = \underset{\sim}{x}_1))}{P(A|\underset{\sim}{X} = \underset{\sim}{x}_2) / (1 - P(A|\underset{\sim}{X} = \underset{\sim}{x}_2))}. \tag{2.9}$$

The numerator of (Eq. (2.9)) is the odds of success for an individual with covariate vector $\underset{\sim}{x}_1$ and the denominator is the odds of success for an individual with covariate **vector** $\underset{\sim}{x}_2$. The $OR(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ is a measure of how much more likely is a positive outcome under $\underset{\sim}{x}_1$ than under $\underset{\sim}{x}_2$.

Risk is the probability of a negative response (i. e., death) given present conditions. If death within an interval is denoted by ***the*** event $\bar{A}$, ***the relative risk*** ($RR(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$) between two individuals with covariate vectors $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$ is:

$$RR(\underset{\sim}{x}_1, \underset{\sim}{x}_2) = \frac{P(\bar{A}|\underset{\sim}{X} = \underset{\sim}{x}_1)}{P(\bar{A}|\underset{\sim}{X} = \underset{\sim}{x}_2)}. \tag{2.10}$$

Relative risk (*RR* (x_1, x_2)) is a measure of how much more likely it is for an animal with covariate vector x_1 to have died in the interval than for an animal with covariate vector x_2.

Another quantity that characterizes *survival* potential *is an* **animal's** *life expectancy.* **If** *T is the* time to death, then the expected lifetime (i.e., $E(T|x)$) of an animal with covariate vector x can be expressed as

$$E(T|x) = \int_0^{\infty} (t \, x|x) \, t \, . \tag{2.11}$$

The instantaneous relative risk **(IRR)** between two individuals with covariate vectors x_1 and x_2, respectively, is the ratio of their hazard functions. For example, under the hazard link, the IRR is:

$$IRR(x_1, x_2) = \frac{\lambda_0(t)\, e^{x'_1 \beta}}{\lambda_0(t)\, e^{x'_2 \beta}} = e^{(x'_1 - x'_2)\beta} \tag{2.12}$$

This quantity is a measure of the differences in risk between two animals on an instantaneous basis.

2.1.3.3 Link functions

Suppose an animal is subject to a set of fixed (time-invariant) covariate values x (either intrinsic traits or exogenous factors) anticipated to influence its survival. When modeling the survival process on the continuous-time scale, it is typical to model the hazard function as a function of *the* covariates: $\lambda(t; x, \beta)$, where β is a vector of effect *parameters.* However, on the discrete-time scale, it is more common to model the survival probabilities in each sampling interval as functions of the covariates: $S_k(x, \beta)$. **In** Program SLJRPH, the covariate relationships *are* assumed to be functions of *the linear predictor* $x'\beta$. The function that connects the linear predictor to the hazard function or **survival** probability *is* known *as the link function. The* most important link functions found in epidemiological literature are listed in Table 2.1.

Table 2.1 **Definitions and properties of potential link functions for modeling capture and survival probabilities as functions of covariates. The parameter θ is a generic capture or survival probabilities.**

Name	Link Function $\alpha + \underline{x}'\underline{\beta}$	Inverse Function[d, g] θ	Hazard Rate[a, e]	Life Expectancy[b, f] (Continuous Time)	Odds Ratio[c]	Response Relationship
Identity	θ	$\alpha + \underline{x}'\underline{\beta}$	N/A	N/A	$\dfrac{(\alpha + \underline{x}'_1\underline{\beta})\,(1 - \alpha - \underline{x}'_2\underline{\beta})}{(\alpha + \underline{x}'_2\underline{\beta})\,(1 - \alpha - \underline{x}'_1\underline{\beta})}$	• additive effect on survival probability
Log	$\log(\theta)$	$\theta e^{\underline{x}\underline{\beta}}$	$\lambda_0(t)\,(\alpha + \underline{x}'\underline{\beta})^{(e)}$	$\dfrac{E[t\|0]}{\alpha + \underline{x}'\underline{\beta}}$	$\dfrac{\theta^{e^{\underline{x}_1\underline{\beta}}}\left(1 - \theta^{e^{\underline{x}_2\underline{\beta}}}\right)}{\theta^{e^{\underline{x}_2\underline{\beta}}}\,1\left(- \theta^{e^{\underline{x}_1\underline{\beta}}}\right)}$	• multiplicative effect on hazard rate • multiplicative effect on survival probability
Complementary Log	$\log(1 - \theta)$	$1 - (1 - \theta)\,e^{\underline{x}'\underline{\beta}}$	N/A	N/A	$\dfrac{\left(1 - (1 - \theta)^{e^{\underline{x}_1\underline{\beta}}}\right)(1 - \theta)^{e^{\underline{x}_2\underline{\beta}}}}{\left(1 - (1 - \theta)^{e^{\underline{x}_2\underline{\beta}}}\right)(1 - \theta)^{e^{\underline{x}_1\underline{\beta}}}}$	• multiplicative effect on mortality probability
Logit or Log Odds	$\log\left(\dfrac{\theta}{1 - \theta}\right)$	$\dfrac{exp\,(\alpha + \underline{x}'\underline{\beta})}{1 + exp\,(a + \underline{x}'\underline{\beta})}$	N/A	N/A	$e^{(\underline{x}'_1 - \underline{x}'_2)\,\underline{\beta}}$	• additive effect on log odds ratio
Hazard or Log-Log	$-\log(-\log(\theta))$	$\theta^{e^{\underline{x}'\underline{\beta}}}$	$\lambda_0(t)\,e^{\underline{x}'\underline{\beta}}$	$\dfrac{E[t\|0]}{e^{\underline{x}'\underline{\beta}}}$	$\dfrac{\theta^{e^{\underline{x}_1\underline{\beta}}}\left(1 - \theta^{e^{\underline{x}_2\underline{\beta}}}\right)}{\theta^{e^{\underline{x}_2\underline{\beta}}}\left(1 - \theta^{e^{\underline{x}_1\underline{\beta}}}\right)}$	• multiplicative effect on hazard rate • exponential effect on survival probability • multiplicative effect on life

Table 2.1 (cont.) Definitions and properties of potential link functions for modeling capture and survival probabilities as functions of covariates. The parameter θ is a generic capture or survival probabilities.

Name	Link Function $\alpha + \underset{\sim}{x}'\underline{\beta}$	Inverse Function(d, g) θ	Hazard Rate(a, e)	Life Expectancy(b, f) (Continuous Time)	Odds Ratio(c)	Response Relationship
Complementary Log-Log	$\log(-\log(1-\theta))$	$1-(1-\theta)e^{x'\beta}$	N/A	N/A	$\dfrac{\left(1-(1-\theta)^{e^{x_1\beta}}\right)(1-\theta)^{e^{x_2\beta}}}{\left(1-(1-\theta)^{e^{x_2\beta}}\right)(1-\theta)^{e^{x_1\beta}}}$	• exponential effect on mortality probability

(a) Hazard rate is the instantaneous rate of mortality.

(b) Assuming constant hazard rate and time-invariant covariates $\underset{\sim}{x}'$.

(c) Measures relative odds of survival between individuals with covariates $\underset{\sim}{x}'_1$ and $\underset{\sim}{x}'_2$, respectively.

(d) Where $\theta = e^{\alpha}$ survival or capture probability with covariates equal to 0.

(e) Where $\lambda_0(r)$ is the hazard rate with covariates equal to 0.

(f) Where $E[t \mid 0]$ is the expected life span with covariates equal to 0.

(g) Where $\theta = e^{\left(-e^{\alpha}\right)}$ survival or capture probability with covariates equal to 0.

An important class of ***link functions are theproportional hazards* (PH) *functions*** **(Feigl** and Zelen 1965; Kalbfleisch and Prentice 1980, pp. 32-33). Program SURPH derives its name as an acronym for **"SURvival** with **Proportional Hazards."** However, other link functions may also be used in Program **SURPH**.

To describe the proportional hazard model, let $\lambda(t;\underset{\sim}{x})$ represent the hazard function at time t for an individual ***with*** covariates $\underset{\sim}{x}$. Initially, the covariates are assumed to be measured at the beginning of the study and to remain constant throughout the study. Under the PH model implemented in SURPH, we have

$$\lambda(t;\underset{\sim}{x}) = \lambda_0(t)\, e^{\underset{\sim}{x}'\underset{\sim}{\beta}}, \tag{2.13}$$

where $\lambda_0(t)$ is an arbitrary, unspecified baseline hazard function. That is, $\lambda_0(t)$ is the hazard function for an individual with covariate vector equal to $\underset{\sim}{0}\ (\underset{\sim}{x} = \underset{\sim}{0})$, or "baseline hazard." The covariates have a multiplicative effect on the hazard function, so that individuals with different covariates have proportional hazard functions. Provided the covariates do not change in value through time, the ratio of the hazard functions **will** also be constant through time. **If** $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$ are vectors of covariates for two different individuals, the ratio of hazard functions can be written as:

$$\frac{\lambda(t;\underset{\sim}{x}_1)}{\lambda(t;\underset{\sim}{x}_2)} = \frac{\lambda_0(t)\, e^{\underset{\sim}{x}'_1\underset{\sim}{\beta}}}{\lambda_0(t)\, e^{\underset{\sim}{x}'_2\underset{\sim}{\beta}}} = e^{(\underset{\sim}{x}'_1 - \underset{\sim}{x}'_2)\underset{\sim}{\beta}} \tag{2.14}$$

Now, consider the survival probabilities under discrete-time sampling for an animal subject to a vector of covariates. The conditional survival probability for a particular period i is the baseline probability raised to the power $e^{\underset{\sim}{x}'\underset{\sim}{\beta}}$, where

$$S_k(\underset{\sim}{x}) = exp\left(- \int_{t_{k-1}}^{t_k} \lambda_0(u)\, e^{\underset{\sim}{x}'\underset{\sim}{\beta}} du \right) = S_k^{e^{\underset{\sim}{x}'\underset{\sim}{\beta}}}. \tag{2.15}$$

Under the proportional hazards model (2.9), the link function implies that the survival probability is related to the linear predictor by a "log-log" transformation where

$$\ln(\ln S_k) = \underset{\sim}{x}'\underset{\sim}{\beta}. \tag{2.16}$$

Thus, this link function is sometimes called the "log-log link," though we will use the term "hazard link."

A second common link function is the logistic parameterization (Cox and Snell 1989):

$$S_k(\underset{\sim}{x}) = \frac{exp(\underset{\sim}{x}'\underset{\sim}{\beta})}{1 + exp(\underset{\sim}{x}'\underset{\sim}{\beta})}, \tag{2.17}$$

a common alternative for modeling survival probabilities. The logit-link function

$$ln\left(\frac{S}{1-S}\right) = \underset{\sim}{x}'\underset{\sim}{\beta} \tag{2.18}$$

relates the linear predictor to a nonlinear function of survival. Like the hazard link, the logit link is popular because of its mathematical property that bounds survival estimates within the admissable range of O-l. Lebreton et al. (1992) extensively used the logit-link function in their monograph on tag analysis.

Lebreton et al. (1992) also describe the linear-link function for modeling survival where

$$S(\underset{\sim}{x}) = \underset{\sim}{x}'\underset{\sim}{\beta}. \tag{2.19}$$

However, a mechanistic argument for a linear-link function is generally absent. Instead, the choice of (2.19) is reminiscent of standard linear regression models often used in the absence of a well-defined response model.

A final important survival link function often used in epidemiology is the log link where

$$ln(S(\underset{\sim}{x})) = \underset{\sim}{x}'\underset{\sim}{\beta}. \tag{2.20}$$

derived from the survival function

$$S(\underset{\sim}{x}') = S_0 e^{\underset{\sim}{x}'\underset{\sim}{\beta}}, \tag{2.21}$$

The log link also gives a proportional hazards model; however, the effects on survival probabilities are multiplicative rather than exponential in nature.

For the four link functions for survival discussed so far (i.e., 2.16, 2.18, 2.19, and 2.20), complementary links can be established by modeling mortality (i.e., 1-S) rather than survival (S). Hence, a suite of possible response models are available to investigators. Program SURPH currently provides the two most important link functions for survival and capture probabilities; the hazard (log-log) and logit.

The selection of a particular link function is based on several considerations, including the nature of the process that gives rise to the survival probability, the response relationship implied, and the ease of calculating summary statistics (i.e., 2.9, 2.10, 2.11, 2.12). Both the log and hazard links have been used to describe continuous-time survival processes (Kalbfleisch and Prentice 1980), while the logit link is often used when the data are binomial counts of successes and failures (i.e., binomial outcomes of discrete events) (Hosmer and Lemeshow 1989). IRR and $E(T)$ can be calculated from parameters under hazard link but not under the logit link. The hazard function is a description of instantaneous risk through time, whereas the logistic parameterization can only give a description of risk for an entire interval. Quantities useful for describing continuous processes, such as the instantaneous hazard rate and life expectancy, cannot be calculated from the logit link but are easily derived using the log- and hazard-link functions. The odds ratio is more easily calculated from the parameters under the logit link.

The alternative link functions (Table 2.1) can have dramatic differences in the shape of the response curve (Figure 2.1). The curves in Figure 2.1 are constrained to pass through the points of 100% mortality and 100% survival at the low and high end of the covariate gradient. These constraints maximize differences among the various link functions. When the range of survival probabilities is restricted, the differences among the link functions are much smaller. For example, Figure 2.2 shows a case where the survival functions are forced through the points of 70% and 100% survival. Here the logit, hazard, and complementary log functions are nearly

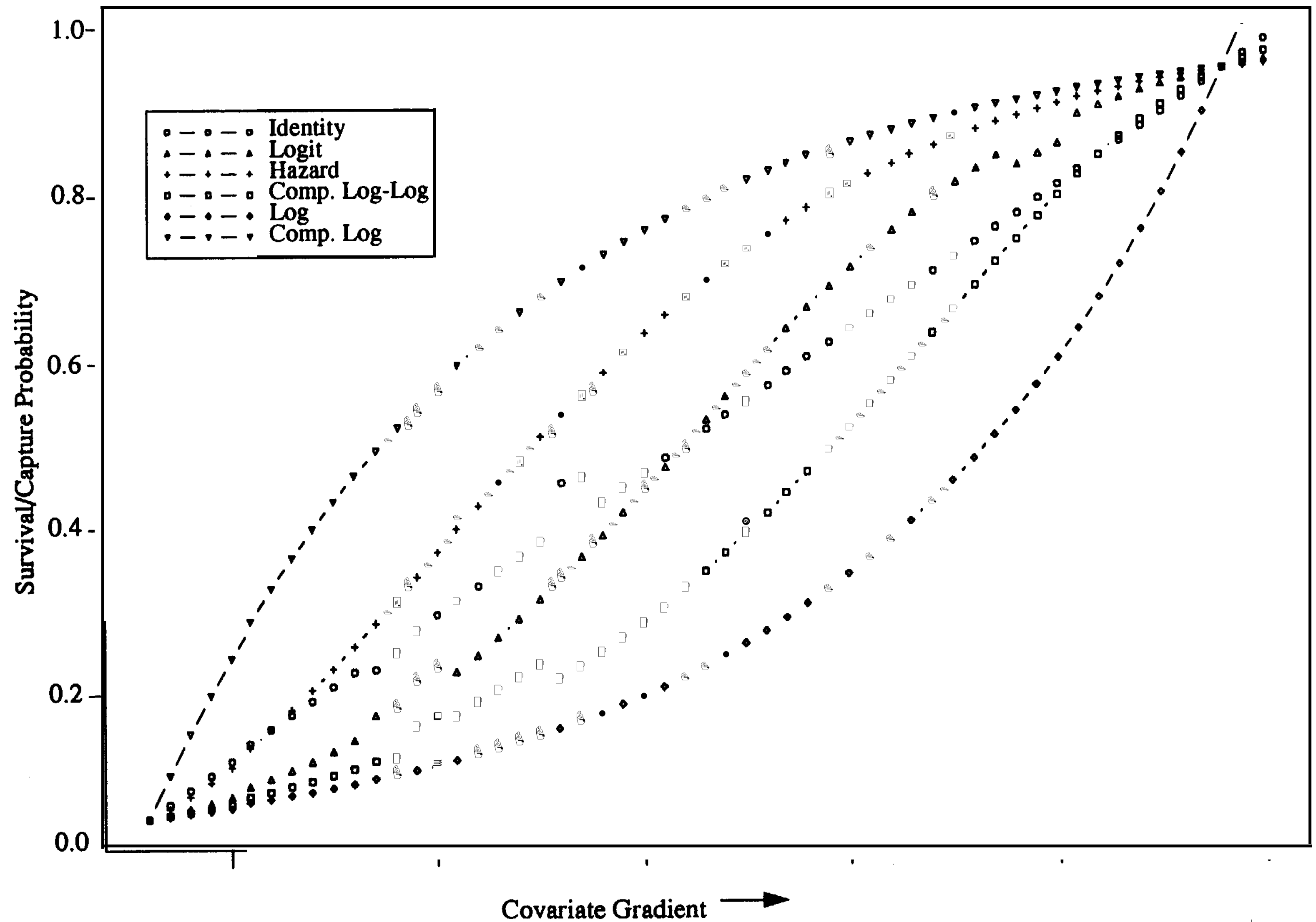


Figure 2.1 Graphical illustration of link functions for modeling relationships between covariates and capture or survival probabilities. All functions go through 0.0 at low end of covariate gradient and 1.0 at high end.

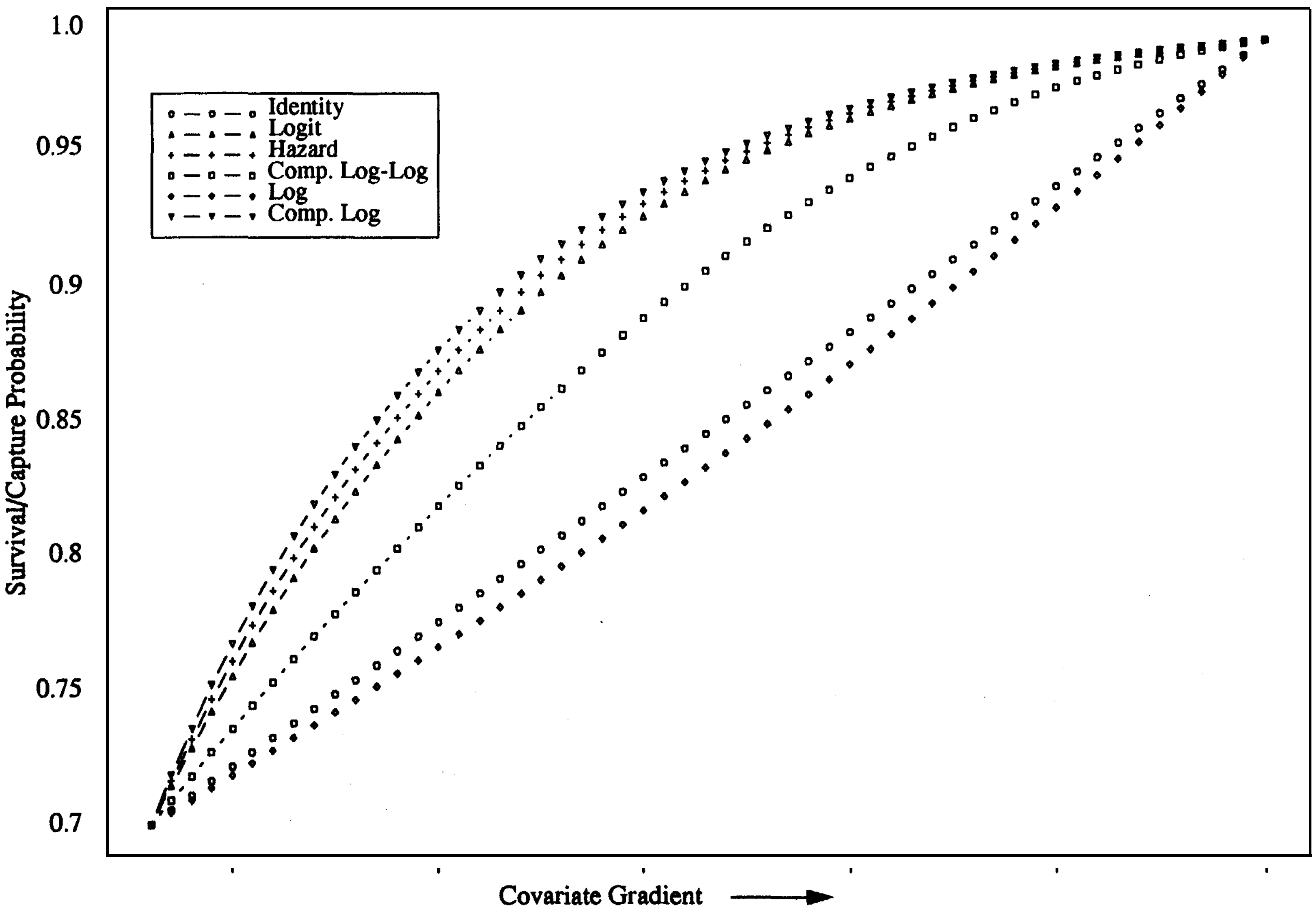


Figure 2.2 **Graphical illustration of link functions for modeling relationships between covariates and capture or survival probabilities. Range of probabilities constricted to 0.7 to 1.0.**

indiscernible within the restricted range of 0.7 and 1.0. In **Figure** 2.3, probabilities range from 0.0 to 0.3, and the **logit,** complementary log-log and log links all produce similar curves in the range. Outside the restricted ranges, the various link functions can vary widely. Consequently, if the estimated survival probabilities for a particular data set fall in a restricted range, extrapolation outside the range is not recommended.

In actual tag-release studies, the range of survival probabilities is usually relatively narrow. The similarity of link functions in restricted ranges **implies** that the choice of link functions in many analyses may not be critical with regard to fitted (predicted) survival probabilities and the resulting maximized likelihood value. Hence, inferences regarding tests of effects on survival (using likelihood ratio tests or analysis of deviance) will almost always be the same regardless of the link function employed. The major difference among link functions, then, is the interpretation of the parameters, and the implied effects on the hazard rate, life **expectancy,** and relative risk.

Instantaneous relative risk **(Eq.** (2.21)) and life expectancy **(Eq.** (2.20)) can be calculated from the parameters of the hazard link but not under the **logit** link. The hazard function is a description of instantaneous risk through time, whereas the logistic **parameterization** gives a description of risk for the entire interval. If environmental conditions are constant (i.e., λ_0 (t) = A,), the life expectancy of an animal with covariate vector $\underset{\sim}{x}$ under the hazard link is:

$$E(T|\underset{\sim}{x}) = \frac{1}{\lambda_0 e^{\underset{\sim}{x}'\underset{\sim}{\beta}}} = \frac{E(T|0)}{e^{\underset{\sim}{x}'\underset{\sim}{\beta}}} . \tag{2.22}$$

Thus, covariates have a multiplicative effect on life expectancy under the hazard link.

The probabilities of capture can also be modeled as functions of individual- and **group-** covariates. Although any of the link functions in Table 2.1 can be used, we recommend use of the **logit** link because capture probabilities in release-recapture models represent success probabilities for discrete-time events (the probability of being captured on a particular sampling occasion). Program SURPH can model capture probabilities in release-recapture studies using either the **logit** or hazard link.

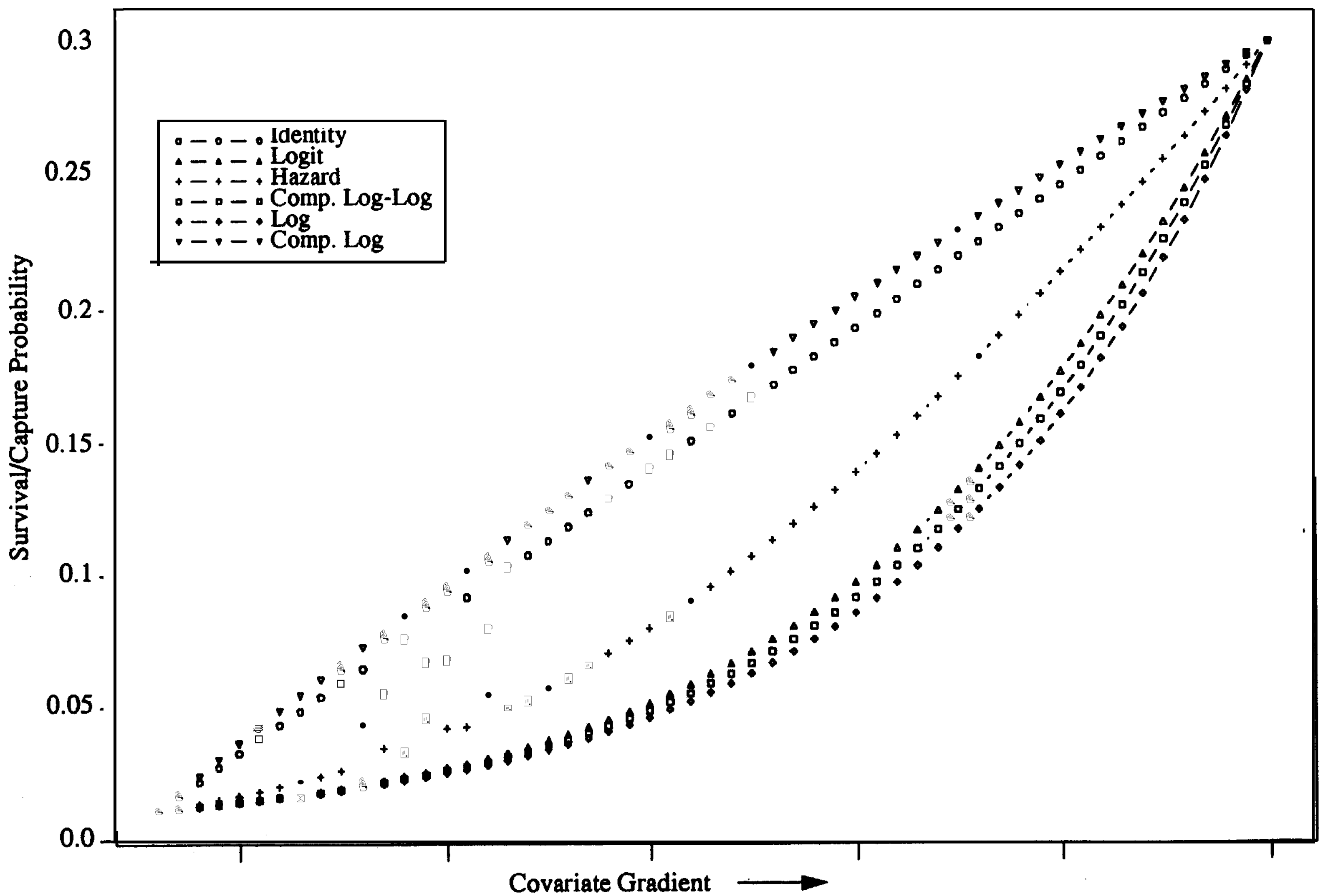


Figure 2.3 **Graphical illustration of link functions for modeling relationships between covariates and capture and survival probabilities. Range of probabilities constricted to 0.0 to 0.3.**

2.2 Design of Tag-Release Survival Studies

The ability to detect and quantify effects of covariates on survival rates is a powerful tool for fish and wildlife managers, regulatory agencies, population biologists, and others. Effects on survival may have both spatial and temporal aspects. Researchers might be interested in how survival rates differ among populations with different levels of a covariate within a particular time period, or they might wish to learn how survival probabilities at a particular site change through time as the level of a covariate changes. Multiple-site tag-release studies have both spatial and temporal aspects to then. study design and are well-suited to the problem of assessment of effects.

Known-fate and release-recapture studies can be implemented to study survival over time intervals or over distances traveled. Thus, the experimental design can be thought of in one of two frames of reference: time or distance. For species that are faithful to a definable study area, survival of animals over time intervals and selection through time are the primary quantities of interest. Survival and selection over distance traveled are the primary quantities of interest in studies on migratory species where individuals share the same migratory route. Salmon smolt outmigration in the Columbia-Snake River system is an example where the frame of reference is distance. Other examples include studies on migratory birds, such as the Ross goose, where the migratory route is well-defined (Welty 1982). Known-fate and release-recapture studies can be applied equally well to studies on selection whether the frame of reference is time or distance (Schwarz 1988).

The type of covariates and the study designs used to investigate them will depend on the goals of the survival study. For example, agrochemical companies and regulatory agencies will wish to reach unambiguous conclusions regarding the effects of a pesticide on survival rates. To this end, they will conduct carefully controlled, multiple-site experiments according to the tenets of experimental design. The covariate of central interest will probably be the level of pesticide application on each site; a covariate that is defined at the population level. On the other hand, the questions of concern to population biologists, general ecologists, and others may not lend themselves to study by controlled experiments. Indeed, it may be impossible to control the covariates. These researchers may seek general insights into factors influencing population dynamics, and may obtain information from observational studies. Observational or descriptive studies can be used to establish correlative, but not causal, relationships between environmental covariates and survival rates. Covariates for observational studies might include period- or site-specific factors and/or traits measured on individual animals to investigate selective pressures.

In the sense intended here, ***an experimental study*** is one that involves manipulation of the environment in a controlled and prescribed way. Federer (1973) provides a useful definition of an experiment as the collection of measurements or observations according to a prearranged plan for the purposes of obtaining evidence to test a theory or hypothesis. The goal is to isolate the effect of the manipulation on survival rates. In other words, experimental studies attempt to control for all environmental factors except for the factor being manipulated to allow us to demonstrate that changes are caused by the manipulation. ***An observational study,*** on the other hand, is one in which the covariate levels are not under the control of the investigator. The investigator conducts the tag-release study and records how covariates vary over space and tune. Observational studies can be used to show that survival probabilities are correlated with covariates but cannot demonstrate causality because of the absence of experimental control. When an observational study shows correlation between covariates and survival, it can suggest hypotheses to be tested in subsequent experimental studies.

2.2.1 Principles of experimental design

Cox (1958) gives three conditions that prompt an experimental approach to answering a research question: (1) an objective of comparing treatment effects, (2) a substantial variation in response from plot to plot in the absence of a treatment effect, and (3) treatment differences that are relatively stable despite possible fluctuations in mean response levels. Fisher (1947) admonishes that the study must provide a valid estimate of the error variance with which to test hypotheses. This error variance is a measure of the extraneous variation that tends to mask the treatment effects. Fisher (1947) states that the correct measure of the error variance for testing treatment means is “one that contains all sources of variation inherent in the variation among treatment means except that portion due specifically to the treatments themselves.” To assure the validity of the comparison of treatments, an experimental design must adhere to the fundamental principles of replication, randomization, and experimental control.

If the effects of a treatment are to be demonstrated spatially (i.e., sites with different levels of the treatment have different survival rates), it is not sufficient to show that a single site with one level of the treatment differs from another site with a different level. The effects of the treatment in this case are completely confounded with the effects of any other differences between the two sites. The solution is replication, a basic technique ***in*** experimental design. ***Replication is the*** application of similar treatment conditions to two or more experimental units (sites). Replication accomplishes three major objectives: (1) It allows the experimenter to show that the treatment effect is repeatable; (2) it allows estimation of experimental error or error variability; and (3) it leads to a more precise estimate of treatment effects.

A slightly less stringent version of replication is seen in the design of a “controlled regression” study. In controlled regression, the levels of a covariate are under the control of the investigator and applied at random to experimental units. If a particular form is assumed for the relationship between a covariate and survival response, then each level of the covariate need not be exactly replicated. The assumed form of the relationship provides the opportunity to estimate the experimental error using the deviations from the model as a measure of variability. A classic example of a controlled regression study is the ***dose-response*** design, where differing levels of a treatment are applied to the replicate experimental units at random.

Randomization is employed to distribute the idiosyncratic characteristics of the experimental units over the treatment levels so that they will not systematically bias the outcome of the experiment. Thus, randomization provides a basis of obtaining unbiased estimates of the treatment effects in the face of unassigned variability among replicated experimental units. Randomization also helps to ensure that the measured responses are **statistically** independent (Kirk 1982). In this context, “random” does not mean “haphazard”; rather, that treatments are assigned to the experimental units in a probabilistic manner. The experimental design is a plan by which treatments are randomized to sites. All familiar designs for linear models involve randomization, including randomized block, Latin square, split plot and the completely randomized design. All of these designs can be analyzed using Program SURPH.

The use of ***experimental controls*** implies that experimental units under treatment conditions are compared with units ("controls") treated identically except for the treatment under scrutiny. Often, treatment conditions will be compared to "null" or "normal" conditions. The null condition is often the complete absence of the treatment. In agricultural tests of a pesticide, for example, the control sites would not receive an application of the pesticide, though they may have an inert substance similar to the pesticide sprayed to control for the direct physical effects of the effort of applying the pesticide. Alternatively, the control condition may be an established level of an environmental factor for comparison with elevated or depressed levels of the factor.

Using the three principles outlined above, the choice of experimental design (i.e., the restrictions used in randomizing treatments to site) will depend on the environmental factors one wishes to control through blocking and the constraints imposed by the location and number of available sites. Typically, the simplest design that adequately controls for extraneous sources of variability should be selected. For simple comparisons of two treatments, completely randomized or paired designs will be adequate more often than not.

2.2.2 Release and sampling protocols

The basic elements of an animal survival study are the same for known-fate and **release**-recapture studies. The protocol consists of two components: (1) the designation of sampling events upon which survival information is obtained; and (2) a release design that specifies how newly tagged animals enter the study. The discussion below includes cases when the frame of reference is time or distance.

The first component is the designation of the sampling events. Sampling events are those occasions in time or those locations in space where survival status is determined in known-fate studies or where recapture efforts are made in release-recapture studies. The sampling events divide the study into intervals. The term "interval" is generic and refers to the time or distance between sampling events. When time is the frame of reference, the intervals are called periods. When distance is the frame of reference, the intervals are segments. For the special case of salmon outmigration studies, the segments are river reaches. Figure 2.4 illustrates the numbering scheme for sampling events and intervals.

The sampling events are "snapshots," giving a survivor profile of the marked animals at a point in time or space. The sampling events should be close enough in time or space that external environmental conditions within each interval vary as little as possible. If the conditions within an interval are not constant, individuals with one set of traits may be favored in one part of the interval and those with a different set of traits favored in a different part. In this case, the effects may cancel out, so that over the entire interval, no selection can be detected even though it occurred (Endler 1986, Chapter 4). Constancy of environmental conditions over an interval increases the chances of detecting and correctly interpreting selection if it exists. In addition, if the external conditions are to be used as group-based covariates in a multi-site study, volatile conditions within intervals makes it difficult to characterize the conditions in a meaningful way.

Figure 2.4, with sampling events depicted as single points along the line, illustrates the assumption that sampling events occur essentially instantaneously. That is, it is assumed that sampling events occur over a very short period of time or distance relative to the length of the intervals between events. Instantaneous sampling assures that the intervals over which survival is estimated are well-defined. If the events are not instantaneous (i.e., sampling occurs over a large span of time or space relative to the length of the intervals between sampling events), the beginning and end points of the intervals become "fuzzy." A consequence of non-instantaneous sampling is illustrated in Figure 2.5. The figure illustrates the fate of two animals released at the same time in a known-fate study. Animal B actually survives for a longer time than A, yet A would be recorded as having survived the first interval and B would be recorded as not having survived. The probability of survival from release until the first sampling event is not a well-defined quantity when significant mortality can occur during a sampling event.

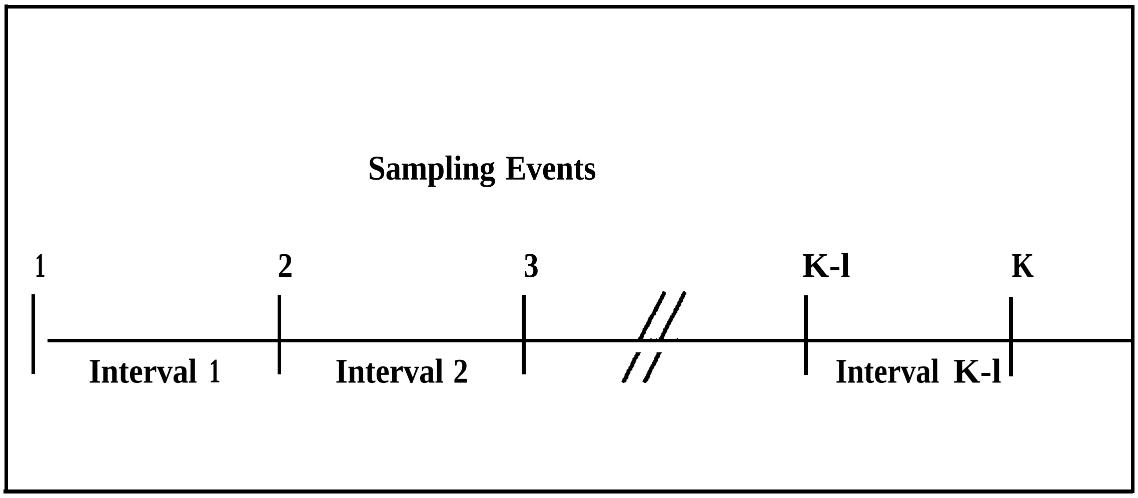


Figure 2.4 Schematic of the numbering system for the sampling events and intervals. The K sampling events divide the study into K-l intervals over which survival is studied.

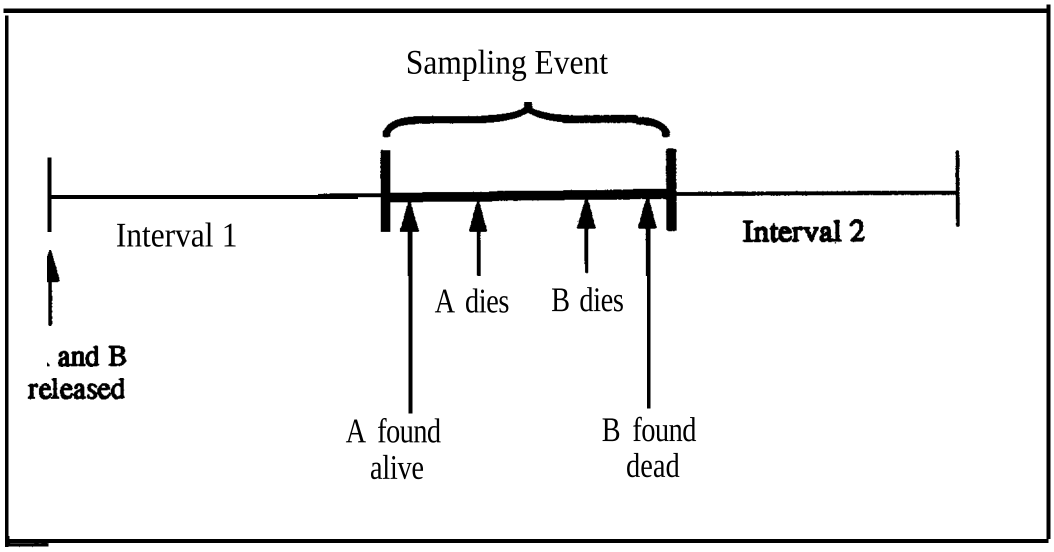


Figure 2.5 Illustration of a problem that can occur when mortalities can occur during a prolonged sampling event.

When the frame of reference is time, the problem is to **find** all the animals in the shortest time possible. The duration of a sampling event will depend on the techniques used for marking and recapturing the animals and on the number of personnel available. When the frame of reference is distance, fixed observation posts positioned at points along the migratory route assures that the instantaneous sampling assumption will be satisfied. It is important that the observation posts be operational until every survivor has passed the point.

The second component of the study protocol is the release design, which determines how tagged animals enter into the study. The release design consists of the number of populations included in the study and by the number of release groups per population. In this manual, for purposes of tag-release studies, the term "population" has a very specific meaning, not necessarily coinciding with the usual ecological definition. A population is a distinct group of animals to which the same levels of all group covariates apply, or which share common traits, such as sex or age class. For example, in a multi-site study, each site may be assumed to have a single, distinct population, with group covariates characterizing the sites. If survival is expected to differ between sexes, the analysis could include two populations on each site, defining males and females as separate populations (note that sex could also be defined as an individual-based indicator variable). In a study with the distance frame of reference, such as a juvenile salmon study on the Columbia River, populations might be defined by the tune of release. For example, if releases are made at the same location on three separate occasions, the fish in each release are considered to be distinct populations.

A "release group" is defined as the collection of animals within a population marked and released at a particular sampling event. This definition differs from that used by **Burnham** et al. (1987) where a release group is defined to be the collection of animals captured and released at a particular sampling event, including previously marked individuals. Under the definition of **Burnham** et al. **(1987),** an animal is a member of as many release groups as the number of sampling events on **which** it was captured and released. Defining release groups according to the event when the animals were first captured and marked is convenient for our purposes, because animals are grouped by the event at which their individual covariates were measured. When the frame of reference is tune, the release groups are **defined** by the time when the animals entered into the study. When the frame of reference is distance, the releases are defined by the location where the animals entered the study. Accordingly, each animal is a member of only one release group, regardless of subsequent capture history. Within a release group, the covariates measured on each animal are comparable because they are measured at the same time or at the same place.

There are four basic study designs defined by the number of populations involved and the number of release groups within each population: (1) single-population, single-release design; (2) single-population, multiple-release design; (3) multiple-population, single-release design; and (4) multiple-population, multiple-release design.

2.2.2 .I Single-population single-release

In the single-population, single-release design **(Figure 2.6),** a single random sample of ***R*** animals is obtained from a designated biological population. *This* is *also known as* a *single-entry* release design. All individuals are measured and released into the study area simultaneously at Event 1. With this design, a survival-estimation model can be used to obtain estimates for interval-specific survival probabilities S,, S,, . . ., S_{K-1}. With a survival-effects model, one can study relationships between the survival probability in each interval and group-based covariates specific to each interval. In addition, the relationship between survival in each interval and individual traits can be examined. The survival parameters shown in Figure 2.6 are defined as follows:

S = baseline, or intercept parameter for survival probabilities;

ρ_k = regression parameter describing the difference from the baseline survival probability for interval k (k= 2,..., K-l);

β_{kl} = regression parameter for the relationship between the Zth (l = 1,..., L) **individual-**based covariate and survival probability in the k**th** interval (k= 1,..., K-1). **These** parameters are also referred to as selection parameters; and

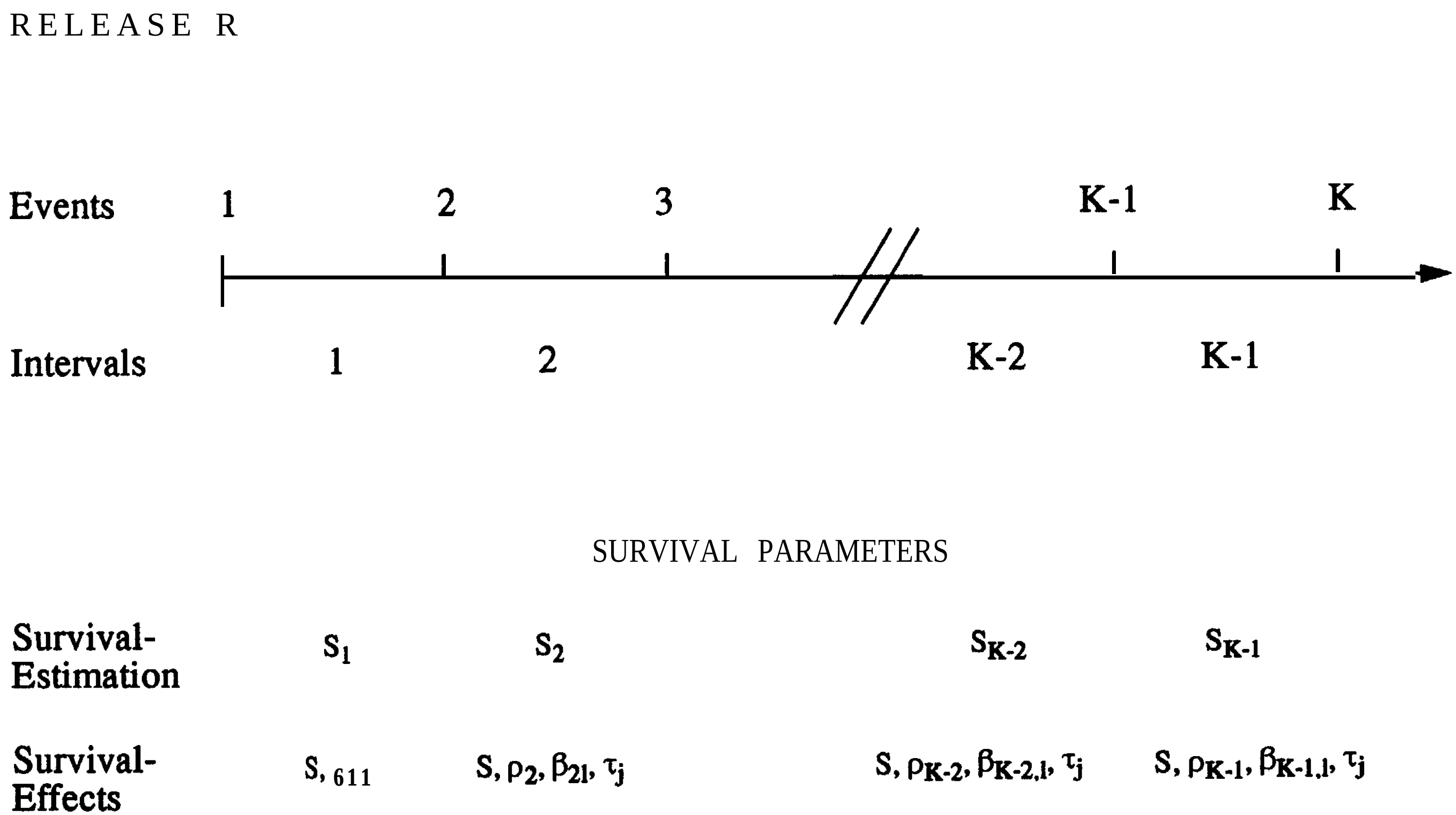


Figure 2.6 **Schematic of single-population, single-release design, showing parameters that can be estimated using survival-estimation and survival-effects models. Design consists of a single release of *b* uniquely tagged animals at event 1.**

τ_j = regression parameter for the relationship between survival probabilities and the jth interval-specific environmental (group-based) covariate (j=1,..., J).

Using release-recapture data, it is also necessary to model the capture probabilities. In a capture-estimation model, the event-specific capture probabilities are denoted $P_2, P_3, \ldots, P_K$. The survival probability in the final interval and the capture probability on the final sampling event cannot be modeled separately. Release-recapture data have this restriction because it is impossible to know whether an animal not seen on the final sampling event died during the last interval or if it survived and eluded detection. It is possible to model the product of the probability of survival through the final interval and probability of detection at the final event ($S_{K-1}P_K$) (also known as the "product term"). In addition to the survival-process parameters above, the following parameters can be estimated using release-recapture data:

P = baseline, or intercept parameter for capture probabilities;

η_k = regression parameter describing the difference from the baseline capture probability for sampling event k (k= 3,..., K);

δ_{kl} = regression parameter for the relationship between the lth (l = 1,..., L) individual-based covariate and capture probability on the kth event (k= 2,..., K);

γ_j = regression parameter for the relationship between capture probabilities and the jth event-specific (group-based) covariate (j= 1,..., J).

λ = baseline, or intercept parameter for product term; and

χ_l = regression parameter for the relationship between the lth (l = 1,..., L) individual-based covariate and the product term.

With only a single population, the product term cannot be modeled as a function of group-based covariates.

With the single-population, single-release design, one can test whether the survival process over time is related to external factors that vary with time (i.e., $H_0{:}\rho_k = 0$ or $H_0{:}\tau_j = 0$). In addition, hypotheses about the selection pressures among animals within sampling intervals (i.e., individual-based covariate effects) can be tested (i.e., $H_0{:}\beta_{kl} = 0$). Analogously, changes in capture probabilities over time (i.e., $H_0{:}\eta_k = 0$ or $H_0{:}\gamma_j = 0$) and as a function of individual-based covariates (i.e., $H_0{:}\delta_{kl} = 0$) can be tested.

2.2.2.2 Single-population multiple-release

In the single-population, multiple-release design (Figure 2.7) with K sampling events, there can be as many as K- 1 release groups of size R_k $(k = 1,..., K - 1)$, one on each sampling event except the last. The individuals in each release group are drawn from the same population. When releases of newly-marked animals occur at multiple sampling events, the release design is called ***staggered-entry.*** Without staggered entry, the number of marked animals surviving until the later stages of the study can be too small to estimate parameters with reasonable precision. The parameters that can be estimated are the same as for the single-population, single-release design. Time-variant individual-based covatiates cannot be studied using the multiple-release design, because the covariates are measured at different times for the various release groups. Time-invariant individual covariates can be modeled, as the time of measurement is irrelevant. As in the single-release design, relationships can be investigated between the survival probability in each interval and group-based covariates specific to each interval.

2.2.2.3 Multiple-population single release

In the multiple-population, single-release design (Figure 2.8), a single release group, R_i $(i = 1,..., I)$, is made at the beginning of the study in each of I populations. Animals from all populations are released on event 1. All releases and subsequent sampling events occur simultaneously (i.e., the same time or same location) in all populations. This design is essentially

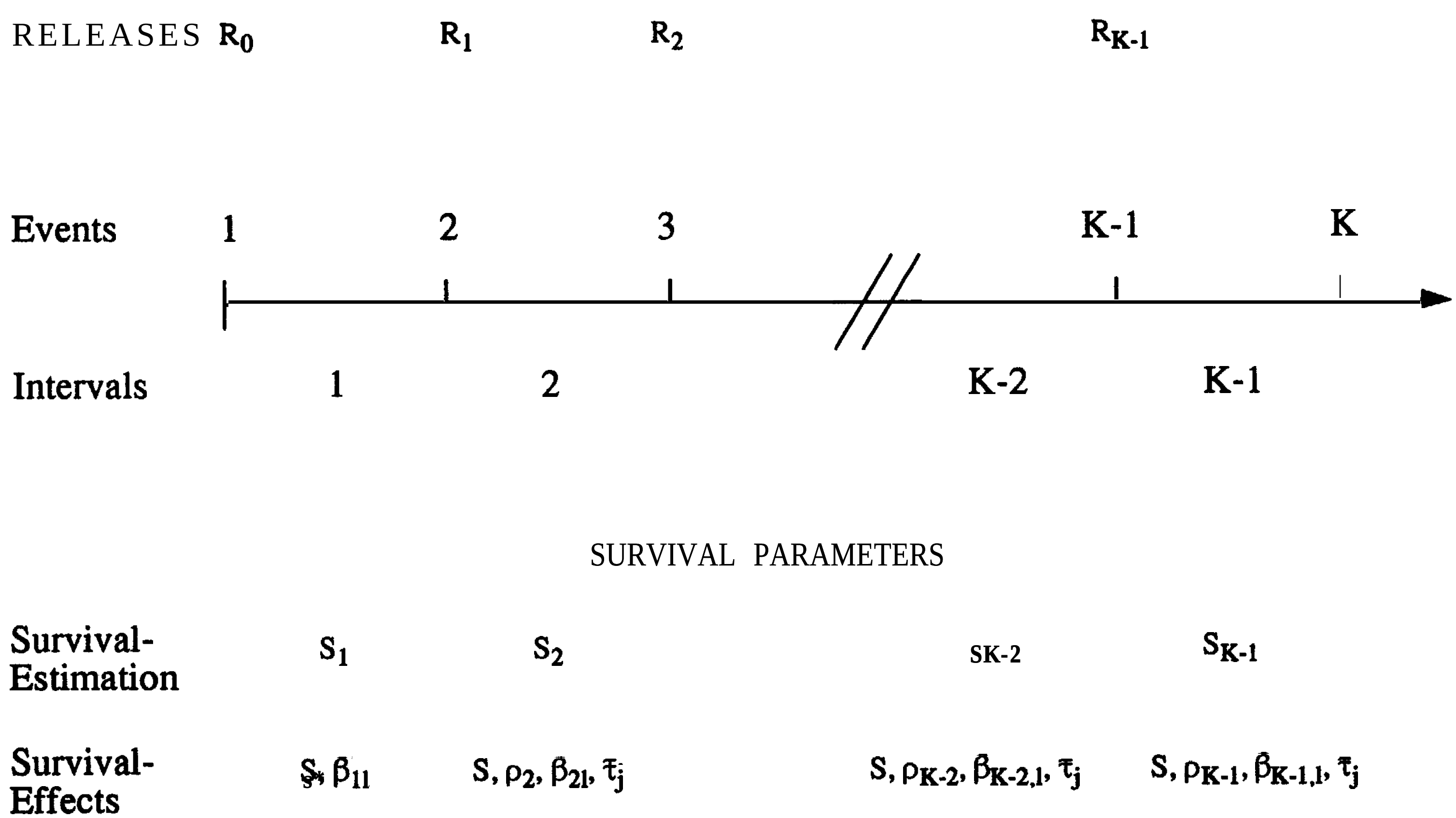


Figure 2.7 **Schematic of single-population, multiple-release design, showing parameters that can be estimated using survival-estimation and survival-effects models. Design consists of as many as K-1 release groups $\{b_k\ (k = 1,\ldots, K-1)\}$ of tagged animals. All animals are drawn from the same population.**

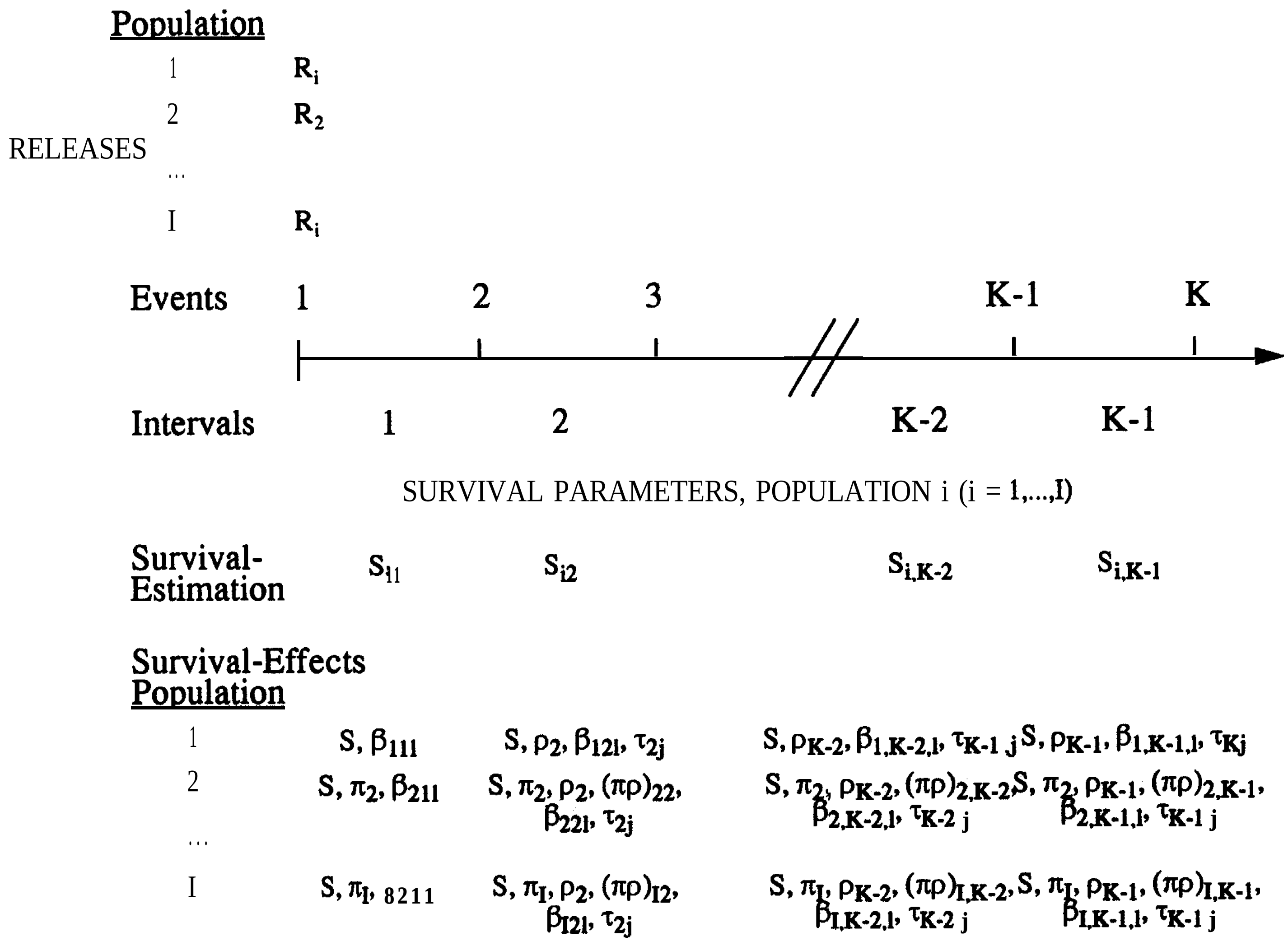


Figure 2.8 Schematic of multiple-population, single release design, showing parameters that can be estimated using survival-estimation and survival-effects models. Design consists of a single release, b_i ($i = 1,..., I$), of tagged animals in each of I populations. The sampling events occur at the same time or location for each population.

the single-population, single-release design carried out on two or more populations simultaneously. If the populations are **defined** by geographic location and environmental conditions vary among the locations, the effect on overall population survival rates of the environmental conditions (group-based covariates) can be modeled. Moreover, interactions between individual-based traits and environmental conditions can be investigated, shedding **light** on selection processes. When segments of a migration path are the focus of the study, the replicate populations are defined by the time of release. The study design can then be used to relate changes in survival probabilities to segment-specific conditions that change through time.

Using survival-estimation models, population- and interval-specific survival probabilities (S_{ik}, $i = 1,...,I$; $k=1,...,K-1$) can be estimated using a multiple-population design. Under a **survival**-effects model, the survival parameters that can be estimated with known-fate data are as follows (Figure 2.8):

S = baseline, or intercept parameter for survival probabilities;

π_i = regression parameter describing the difference from the baseline survival probability for population i ($i = 2,..., I$);

ρ_k = regression parameter describing the difference from the baseline survival probability for interval k ($k = 2,..., K-1$);

$(\pi\rho)_{ik}$ = regression parameter describing the interaction between population and interval effects on baseline survival probability ($i = 2,..., I$) and ($k = 2,..., K-1$);

β_{ikl} = regression parameter describing the relationship between the lth ($i = 1,..., L$) individual-based covariate and survival probability in the kth interval ($k= 1,..., K-1$) for the ith population ($i = 1,..., I$); and

τ_{kj} = regression parameter describing the relationship between survival probabilities in the kth interval ($k= 1,..., K-1$) and the jth ($j = 1,..., J$) environmental covariate measured for all populations in the kth interval.

The following parameters can be estimated using release-recapture data, in addition to the survival-related parameters defined above:

P = baseline, or intercept parameter for capture probabilities;

ν_i = regression parameter describing the difference from the baseline capture probability for population i ($i = 2,\ldots, I$);

η_k = regression parameter describing the difference from the baseline capture probability for event k ($k = 3,\ldots, K$);

$(\nu\eta)_{ik}$ = regression parameter describing the interaction between population and event effects on baseline capture probability ($i = 2,\ldots, I$) and ($k = 3,\ldots, K$);

δ_{ikl} = regression parameter describing the relationship between the lth ($i = 1,\ldots, L$) individual-based covariate and capture probability on the kth event ($k = 2,\ldots, K$) for the ith population ($i = 1,\ldots, I$);

γ_{kj} = regression parameter describing the relationship between capture probabilities on the kth event ($k = 2,\ldots, K$) and the jth ($j = 1,\ldots, J$) environmental covariate measured for all populations on the kth event;

λ = baseline, or intercept parameter for product term (product of final-interval survival and final-event capture probabilities);

φ_i = regression parameter describing the difference from the baseline product term for population i ($i = 2,\ldots, I$);

χ_{il} = regression parameter for the relationship between the lth ($i = 1,\ldots, L$) individual-based covariate and product term for the *ith* population ($i = 1,\ldots, I$); and

ω_j = regression parameter describing the relationship between product terms and the *j*th ($j = 1,\ldots, J$) population-specific environmental covariate measured on the *K*th (final) event.

Multiple-population designs allow significance testing for effects on survival probabilities of group-based covariates that vary among populations, in addition to those that vary across intervals. For individual-based covariates, one can test whether the regression parameters are equal across populations or intervals, i.e., whether or not selection occurs with the same intensity or in the same direction under the different environmental conditions experienced by each population in the study. Similarly, capture probabilities can be modeled as a function of population-specific or event-specific effects and their interactions.

2.2.2.4 Multiple-population multiple-release

The multiple-population multiple-release design (Figure 2.9) combines the aforementioned advantages of both multiple-population and multiple-release designs. That is, capture and survival probabilities can be modeled as functions of population-specific as well as interval-specific covariates, and sufficient numbers of animals are at risk in each interval for precise estimation of survival and capture probabilities.

2.3 Parameter Estimation

Whenever random sampling is used to gather data, the observed data and the values of any statistics used to summarize the data are realizations of random variables (Hoe1 et al. 1971). Statistics derived from random variables, have associated probability distributions. Probability distributions depend on a set of one or more unknown parameters. The parameters of interest here include survival and capture probabilities and regression coefficients that relate survival and capture to covariates. In survival studies, the random variables are the capture histories of the tagged animals. In some models, for instance, the data are summarized by obtaining a count of

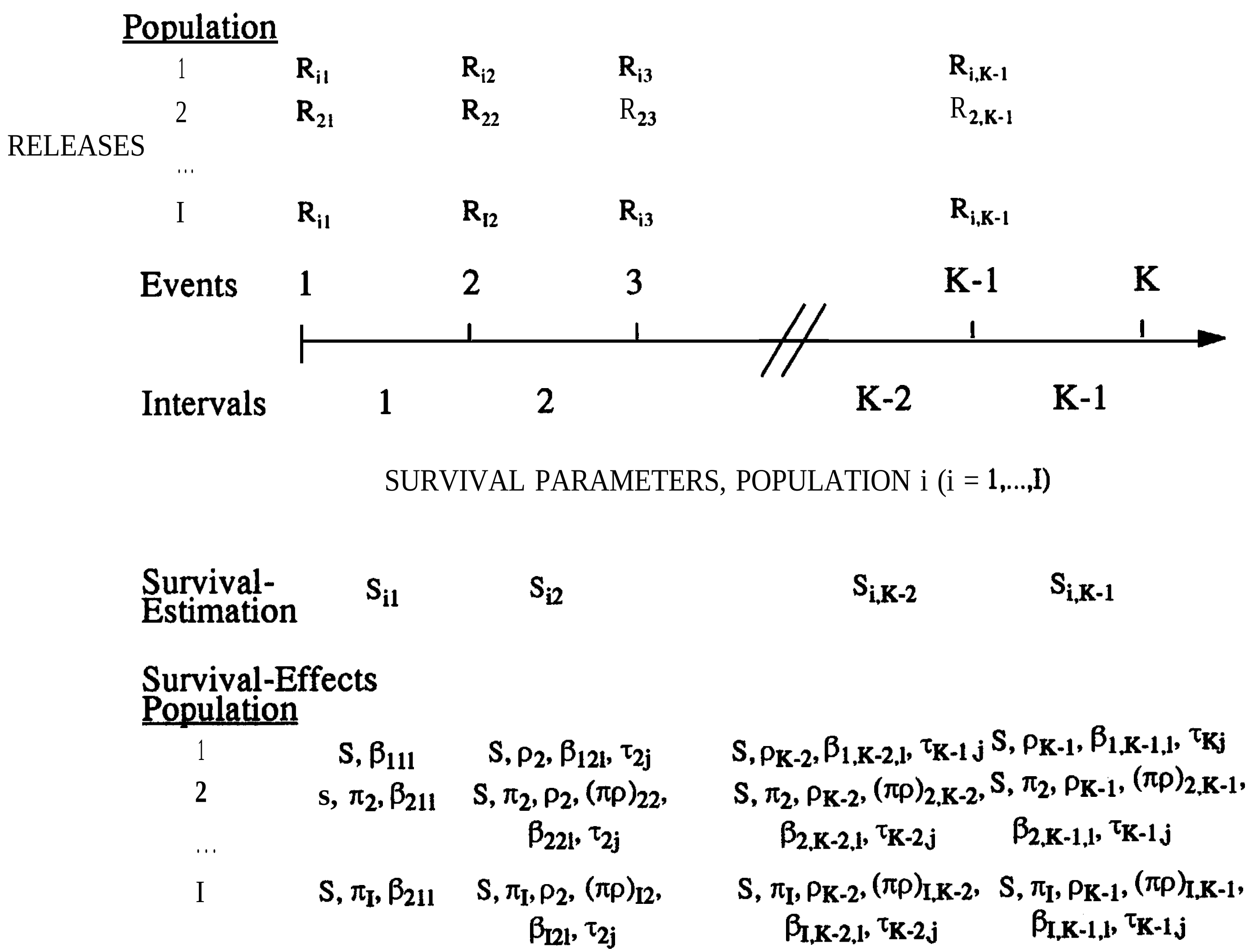


Figure 2.9 Schematic of multiple-population, multiple-release design, showing parameters that can be estimated using survival-estimation and survival-effects models. Design consists of as many as K-1 release groups $\{b_{ik}$ $(i = 1,...,I;$ $k = 1,...,$ K-$1)\}$ of uniquely-tagged animals in each of Z populations.

animals for each of the possible histories. Under certain assumptions, these "cell counts" can be modeled as a sample from a multinomial distribution.

The problem of estimating unknown parameters of a probability distribution from sample data has long been a central concern in the field of statistics, and many methods have been proposed to solve the problem. One of the most well-known and widely-used is the method of ***maximum likelihood (ML) estimation,*** which was presented by Fisher **(1922, 1925)**. The ML method provides a powerful means of deriving point estimates and, equally importantly, estimators of the sampling variances and covariances of the point estimators. Hypothesis testing and model building can also be performed using likelihood theory. The ML method is appealing for its conceptual justification and for its many large sample statistical properties. Maximum likelihood estimators are asymptotically unbiased, normally distributed, and maximally efficient **(Wilks** 1962, Rao 1973, Lehmann 1982).

2.3.1 Maximum likelihood theory

Suppose we have a set of discrete random variables $X_1, X_2, \ldots, X_n$ with joint probability function $f(\underset{\sim}{x};\underset{\sim}{\theta})$ where $\underset{\sim}{\theta}$ is a vector of parameters. In tag-release models, the random variables are usually a set of summary statistics derived from the counts of capture histories (multinomial cell counts). The joint probability function is based on the **multinomial** probability function with cell probabilities modeled as functions of survival and capture probabilities and regression coefficients for covariate effects. When treated as a function of the data, $f(\underset{\sim}{x};\underset{\sim}{\theta})$ gives the probability of the observed data, given the parameter values $\underset{\sim}{\theta}$.

When treated as a function of unknown parameters, on the other hand, the function $f(\underset{\sim}{x};\underset{\sim}{\theta})$ gives the "likelihood" of a set of parameters $\underset{\sim}{\theta}$ given the observed sample data $\underset{\sim}{x}$. The function is then ***called the likelihoodfunction*** for $\underset{\sim}{\theta}$. To emphasize that the likelihood function is considered to be a function of the parameters, it is customary to use the notation $L\,(8.x)$, or simply $L(\underset{\sim}{\theta})$, for the likelihood function. The objective of the ML method is to **find** the set of parameter values that maximizes the likelihood function for a given sample outcome. That is, the estimates are the parameter values under which the sample data are "**most** likely." The parameter values that maximize ***the likelihood*** function ***are known as maximum likelihood estimutes*** **(MLEs)** and are denoted by $\hat{\underset{\sim}{\theta}}$.

Because likelihood functions are often formed as the product of many terms, it is generally easier to deal with the natural logarithm of the likelihood function, or "log-likelihood," which is composed of additive terms. The functions $L(\underset{\sim}{\theta})$ and $\ln L(\underset{\sim}{\theta})$ have their maximum at the same value of $\underset{\sim}{\theta}$. Thus, the **MLEs** $\hat{\underset{\sim}{\theta}}$ are the solution to the following system of equations, known as the "likelihood equations":

$$\frac{\partial}{\partial\theta_1} \ln L(\underset{\sim}{\theta}) = 0$$

$$\frac{\partial}{\partial\theta_2} \ln L(\underset{\sim}{\theta}) = 0$$

$$\vdots$$

$$\frac{\partial}{\partial\theta_v} \ln L(\underset{\sim}{\theta}) = 0$$

where v is the number of parameters to be estimated.

When the likelihood equations can be solved analytically, the parameter estimators $\hat{\underset{\sim}{\theta}}$ are said to be in "closed form." That is, $\hat{\underset{\sim}{\theta}}$ can be written as a set of algebraic functions of the data. When closed forms are not possible (i.e., the likelihood equations cannot be solved analytically) there is a variety of numerical methods designed to find parameter values that maximize the likelihood function. Numerical methods can give estimates when no closed form estimators exist. Numerical optimization methods are discussed in Section 2.5.

As functions of the sample data, the **MLEs** are themselves random variables that **will** vary from sample to sample. Thus, a sampling variance is associated with each estimate. In addition, the estimators $\hat{\underset{\sim}{\theta}}$ are generally correlated, because they are estimated from the same data. Each pair of estimators has an associated sampling covariance. The variances and **covariances** make up ***the covariance matrix,*** which is usually denoted $\underset{\sim}{\Sigma}$:

$$\underset{\sim}{\Sigma} = \begin{vmatrix} Var(\hat{\theta}_1) & Cov(\hat{\theta}_1, \hat{\theta}_2) & Cov(\hat{\theta}_1, \hat{\theta}_3) \ldots & Cov(\hat{\theta}_1, \hat{\theta}_v) \\ Cov(\hat{\theta}_2, \hat{\theta}_1) & Var(\hat{\theta}_2) & Cov(\hat{\theta}_2, \hat{\theta}_3) & Cov(\hat{\theta}_2, \hat{\theta}_v) \\ Cov(\hat{\theta}_3, \hat{\theta}_1) & cov(\hat{\theta}_3, \hat{\theta}_2) & Var(\hat{\theta}_3) & cov(\hat{\theta}_3, \hat{\theta}_v) \\ \vdots & & & \vdots \\ Cov(\hat{\theta}_v, \hat{\theta}_1) & Cov(\hat{\theta}_v, \hat{\theta}_2) & Cov(\hat{\theta}_v, \hat{\theta}_3) \cdots & Var(\hat{\theta}_v) \end{vmatrix}_{v \times v}$$

The matrix is symmetric, since $Cov(\hat{\theta}_i,\hat{\theta}_j) = Cov(\hat{\theta}_j, \hat{\theta}_i)$. **The variances** and **covariances** are usually functions of the sample size and the unknown parameters themselves. Substituting the estimated parameter values of $\hat{\underset{\sim}{\theta}}$ into the expressions for the variances and covariances gives the estimated covariance matrix $\hat{\underset{\sim}{\Sigma}}$. The estimate of $Var(\hat{\theta}_i)$ is denoted $V\hat{a}r(\hat{\theta}_i)$ and the estimate Of $Cov(\hat{\theta}_i, \hat{\theta}_j)$ is denoted $C\hat{o}v(\hat{\theta}_i, \hat{\theta}_j)$.

A property of the ML method is that the asymptotic covariance matrix of the estimators is easily (conceptually, at least) derived from the likelihood function itself. The covariance matrix is the inverse of the expected value of the negative of the matrix of mixed second partial derivatives of the log-likelihood. The matrix of expected values of second partial derivatives is known as the ***Information matrix,*** and is denoted $I(\underset{\sim}{\theta})$ (a function of $\underset{\sim}{\theta}$). The ***ij*th** element of $I(\underset{\sim}{\theta})$ is:

$$-E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln L(\underset{\sim}{\theta})\right].$$

The matrix $I(\underset{\sim}{\theta})$ is evaluated at the true parameter **value** $(\underset{\sim}{\theta})$. If the true values are unknown, the matrix can be evaluated at $\hat{\underset{\sim}{\theta}}$; the **MLEs,** giving the estimated covariance **matrix:** -

$$\hat{\underset{\sim}{\Sigma}} = [I(\hat{\underset{\sim}{\theta}})]^{-1}.$$

The matrix $I(\hat{\underset{\sim}{\theta}})$ is known ***as the observed Information matrix.***

The covariance matrix characterizes the likelihood function in the neighborhood of the maximum point. If the function is very peaked in a particular dimension, then a small movement away from the peak results in a greatly reduced likelihood. In the peaked case, the corresponding parameter can be estimated very precisely, i.e., the variance of the estimator is small. Alternatively, if the function is relatively flat in some dimension, then small movements away from the maximum value make little difference to the likelihood value. In the flat case, the variance is larger; i.e., the parameter is not estimated as precisely with the same sample size.

After variance estimates are obtained, one can rely on the asymptotic normality of **MLEs** to compute confidence intervals. An asymptotic $(1 - a) \times 100\%$ confidence interval for θ is given by:

$$\hat{\theta} \pm Z_{1-\alpha/2}\sqrt{V\hat{a}r(\hat{\theta})}..$$

Asymptotic normality also suggests a statistic that can be used to test hypotheses of the form $H_0\colon \theta = \theta_0$, where θ_0 is a particular hypothesized value for the parameter 8. If $\hat{\theta}$ is normally distributed, then the test statistic:

$$Z = \frac{\hat{\theta} - \theta_0}{\sqrt{V\hat{a}r(\hat{\theta})}}$$

has a standard normal distribution when H_0 is true. To test hypotheses, Z is compared with appropriate quantiles of a standard normal variable (Z-N $(0, 1)$).

Despite its conceptual simplicity, the covariance matrix can be extremely **difficult** to derive, requiring complicated algebra. Most numerical optimization methods (see Section 2.5) use the inverse of the matrix of second partial derivatives, or an approximation to the matrix or its inverse, to find the maximum of the likelihood function. When numerical methods are used, $\hat{\Sigma} = I(\hat{\theta})^{-1}$ can be obtained from the final iteration of the numerical procedure.

2.3.2 Example of maximum likelihood estimation

Consider an experiment to estimate the probability that a win will land "heads" up when flipped into the air. The probability of tossing a "head" is the unknown parameter p , and the probability of a "tails" is (l-p). The experiment wnsists of ***n* trials,** or flips of the win, and the number of heads is recorded, denoted by x.

The number of heads in such an experiment is a random variable with a binomial distribution parameters n (known) and p. The binomial likelihood function for p (given n) is:

$$L(p|n,x) = \binom{n}{x} p^x (1-p)^{n-x} \quad .$$

The log-likelihood can then be written as

$$lnL(p|n,x) = ln \binom{n}{x} + xlnp + (n-x)\,ln(1-p) \quad .$$

The likelihood equation is found by taking the first derivative with respect top, yielding

$$\frac{dlnL(p|(n,x))}{dp} = \frac{x}{p} - \frac{(n-x)}{(1-p)} = 0 \quad .$$

Solving for p in the above equation yields the familiar and reasonable MLE

$$\hat{p} = \frac{x}{n} \quad . \tag{2.23}$$

i.e., the estimate of the true probability of heads is equal to the proportion of heads observed in the sample. The variance of the estimated binomial parameter $\hat{p}$ is

$$Var(\hat{p}) = \frac{p(1-p)}{n} \,. \tag{2.24}$$

The variance is estimated by substituting the parameter estimate (**Eq.** (2.22)) into the variance formula (**Eq.** (2.23)). The variance estimator for $\hat{p}$ can be written as

$$V\hat{a}r(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{n} = \frac{x(n-x)}{n^3}$$

2.3.3 Precision

One of the goals of a survival study may be to estimate a model parameter with a high degree of precision. The other potential goal is to test the significance of a regression coefficient relating survival to measured covariates. Hypothesis tests are discussed in Section 2.4.

A wmmon characterization of precision ***is the coefficient of variation*** (CV) for a parameter estimate, defined as:

$$CV(\hat{\theta}_i) = \frac{\sqrt{Var(\hat{\theta}_i)}}{\hat{\theta}_i} \times 100\% \tag{2.25}$$

The CV provides a rapid means of assessing the relative certainty of a parameter estimate. For example, if the CV is **10%,** then one is approximately 95% certain the estimate does not error by more than **±20%**. This translation of the CV to a measure of error is based on the expression for precision defined by

$$P\left(\left|\frac{\hat{\theta}-\theta}{\theta}\right| < \varepsilon\right) \geq 1-\alpha \,. \tag{2.26}$$

Expression (2.26) specifies that the relative error in estimation (i.e., $(\hat{\theta}-\theta)/\theta$) is less than ε, at least (1 – a) 100% of the time. For example, a precision of **±20%** of the true value of 8, 95% of the time specifies ε = 0.20 and α = 0.05. Skalski and **Robson** (1992, p. 79) show that the probability expression (2.26) can be approximated as

$$P\left(\left|\frac{\hat{\theta}-\theta}{\theta}\right|<\varepsilon\right) \approx 1-2\Phi\left(\frac{100\varepsilon}{CV(\hat{\theta})}\right) = 1 - a \,, \tag{2.27}$$

where CV is expressed as a percentage and Φ devotes the cumulative standard normal distribution. In turn, expression (2.27) solves to the relationship that

$$\varepsilon = Z_{1-\alpha/2}CV(\hat{\theta}).$$

For, α = 0.05, $Z_{1-\alpha/2}$ = 1.96 $\approx$ 2.0, leading to the conclusion that one is approximately 95% certain the error in estimation is less than $\pm 2CV$ (6) . Similarly, one can be approximately 80% certain the error in estimation is less than $\pm 1.3CV$ (8) .

2.4 Hypothesis Testing

In any analysis of tag-release data, a criterion is needed to choose between competing models for describing the data. The process of model building is a series of selections between models. When we test for effects, we are choosing between two models; one that includes the effect of interest and one that does not. Program SURPH uses two basic types of statistical tests of hypothesis. When the inherent variability of the random variables is fully explained by the likelihood model, likelihood ratio tests (LRT) are performed. However, when extra-likelihood variability exists (variability not explicitly explained by the form of the likelihood), analysis of deviance (ANODEV) procedures are used to test hypotheses. Subsequent sections and chapters will describe when these two procedures are appropriate for a particular testing situation.

2.4.1 Likelihood ratio tests (LRT)

The ML method provides a useful, powerful, and general means of selecting between two models when one model is a special case of the other. The test is based on the ratio of the maximized ***likelihoods*** of ***the two*** models and ***is known as*** a ***likelihood ratio test*** (LRT). Two models are “nested,” or “hierarchical,” when one is a special case of the other. For example, Model A is nested in Model B if A can be derived from B by equating two or more parameters of B or by dropping one or more parameters from B. The special case model always has fewer parameters.

When choosing between ***two*** nested models, ***the principle of parsimony dictates*** that the model with more parameters should be selected **if, and** only if, the additional parameters provide a significantly better explanation of the observed data. A model with a larger number of parameters should be chosen only if the observed data is **significantly** more likely under that model. In LRT, the criterion by which nested **models are** judged is the relative values of the maximized likelihoods under each of the models. That is, the model with more parameters is said to provide a better explanation of the data than the other only if its maximized likelihood is significantly greater.

To formalize the discussion, ML estimation is a problem of searching the multidimensional space of possible values of the parameter vector $\underset{\sim}{\theta}$ (the "parameter space") for the point $\hat{\underset{\sim}{\theta}}$ that maximizes the function L $(\underset{\sim}{\theta})$. If we denote the parameter space Θ, then we can define the MLE $\hat{\theta}$ as parameter values such that:

$$L(\hat{\underset{\sim}{\theta}}) = \sup_{\underset{\sim}{\theta} \in \Theta} L(\underset{\sim}{\theta}) .$$

The operator *sup* means "supremum," i.e., find the value of $\underset{\sim}{\theta}$ that maximizes the function L $(\underset{\sim}{\theta})$.

Now, considering the problem of choosing between two nested models, let us refer to the model with more parameters as the "full model," and to the special case model as the "reduced model." If we let the parameter space for the full model be denoted as Θ_F, then the parameter space for the reduced model, denoted Θ_R, is a **subspace** of Θ_F. For example, Θ_R may be the **subspace** where two parameters are equal to each other or where one or more parameters are equal to some hypothesized value(s). Thus, we can write the **maximized likelihoods as**:

$$L(\hat{\underset{\sim}{\theta}}_F) = \sup_{\underset{\sim}{\theta} \in \Theta_F} L(\underset{\sim}{\theta})$$

and

$$L(\hat{\underset{\sim}{\theta}}_R) = \sup_{\underset{\sim}{\theta} \in \Theta_R} L(\underset{\sim}{\theta}),$$

where $\hat{\underset{\sim}{\theta}}_F$ and $\hat{\underset{\sim}{\theta}}_R$ are the **MLEs** under the full and reduced models, respectively. Because Θ_R is a **subspace** of Θ_F, it must be that $L(\hat{\underset{\sim}{\theta}}_R) \le L(\hat{\underset{\sim}{\theta}}_F)$. *Since* strict inequality usually holds, even when the additional parameters are not truly related to the data (the difference is explained wholly by sampling variability in this case), it is necessary to have a criterion to decide what is a significant increase in the likelihood and what can be explained by random sampling variability alone. **LRTs** provide such a criterion.

Likelihood ratio tests are used to test hypotheses of the form

H_0 : Full model is not better than reduced model

versus

H_A : Full model is better than reduced model.

These hypotheses can be rewritten in terms of model parameters. For instance, the hypotheses to test whether survival probabilities in sampling intervals 1 and 2 are equal can be written as

$$H_0: S_1 = S_2$$

versus

$$H_A: S_1 \neq S_2.$$

The full model in this case has one more parameter than the reduced model. The ratio of the maximized likelihoods is:

$$\lambda = \frac{L(\hat{\underset{\sim}{\theta}}_R)}{L(\hat{\underset{\sim}{\theta}}_F)}.$$

The LRT statistic is based on the difference in log-likelihoods:

$$-2\Lambda = -2\,(\ln L\hat{\theta}_R - \ln L(\hat{\theta}_F))\,. \tag{2.28}$$

If the null hypothesis is true, then the statistic $-2ln\lambda$ ***is*** asymptotically distributed as a central **chi-** squared random variable with $V_F - V_R$ **degrees** of freedom, where V_F is the dimension of Θ_F (number of parameters in the full model) and V_R is the dimension of Θ_R (number of parameters in the reduced model). The null hypothesis is rejected if ***-2Znh is*** greater than the quantile of the appropriate χ^2 distribution corresponding to the stated significance level (a) .

2.4.2 Non-hierarchical models

LRTs provide a framework for selecting among hierarchical models. In the course of model selection for any particular data set, however, it is likely that non-nested models will be considered. A simple example of four models for a multiple-population, multiple-interval, **tag-** release study is presented in Table 2.2. The subscripts in the model names refer to the possible dependence of the survival probabilities on space (s) and time ***(t).*** For example, Model S_{st} has survival probabilities unique in space and time, while those in Model S_t depend only on time. The four models define two different hierarchies, namely (S, $\rightarrow S_t \rightarrow$ S) and (S, $\rightarrow S_s \rightarrow$ S) . Models along either hierarchy could be compared using **LRTs**. However, S_t is not nested in S,, and S_s is not nested in S,. Thus, an LRT cannot be used to select between them. A different criterion is needed.

A criterion that has received some attention in tag-release literature (e.g., White 1990) is ***Akaike's Information Criterion (AK) (Akaike*** 1973, Sakamoto et al. 1986). AIC forgoes formal testing by basing selection on the quantity:

$$\text{AIC} = -2\ln L\,(\hat{\underset{\sim}{\theta}}) + 2(\text{number of parameters estimated}). \tag{2.29}$$

The smaller the value of **AIC,** the better the fit of the model to the observed data. Thus, the model selection problem becomes, in effect, one of one-dimensional optimization rather than a multi-dimensional statistical hypothesis testing problem. Lebreton et al. (1992) have advocated the use

Table 2.2 Four alternative models for a multiple-population multiple-interval survival study.

Model	Description
S_{st}	Unique survival probability for each for each interval for each population
S_s	Probabilities population-specific, but constant across intervals
S_t	Probabilities interval-specific, but constant across populations
S	Common survival probability for all populations in all intervals

of only a few formal tests between the model selected under the AIC criterion and a few neighboring ones, thus "limiting the increase in the overall risk of false rejection of at least one null hypothesis otherwise caused by multiple tests." Good results using AIC have been claimed for release-recapture analysis (Lebreton et al. 1992).

The AIC can be used not only for non-nested models, but also when the models of interest are strictly hierarchical. For example, suppose Model $\mathbf{M_1}$ is obtained from $\mathbf{M_0}$ by dropping a single parameter. Referring to **Eq. (2.1)**, we *see* that AIC **will** select $\mathbf{M_0}$ *if* $-2ln\lambda = -2\,(lnL_{M_0} - lnL_{M_1})$ is greater than 2. From the formal testing framework **(LRT)**, we know that *-2Znh* has a χ^2 distribution with one degree of freedom when the null hypothesis **(H_0:M_0** not better than **M_1)** is true. To reject this null hypothesis, when -2Znh is greater than 2.0, implies a **significance** level of 0.16. Questions of multiple testing aside, this procedure has given a risk of Type I error (a) in a single test of 16%. Thus, AIC can be used to identify a set of candidate models from a large number of non-hierarchical models, the criterion should be augmented by formal testing to choose the final model.

2.4.3 Analysis of deviance (ANODEV)

The final important concept in likelihood theory is that of the deviance of a model. The theory of analysis of deviance has largely been developed in the context of generalized linear models, where the data are a set of observed data values from individual study subjects. This context is assumed in the general introductory discussion below. However, analysis of deviance is used in SURPH to assess effects of group covariates, and the "observations" are group- and **interval-**specific survival probabilities, not measurements on individual subjects.

McCullagh and Nelder (1983)'point out that "fitting a model to data may be regarded as a way of replacing a set of observed data values y by a set of fitted values $\underset{\sim}{\hat{\mu}}$ derived from a model involving (usually) a smaller number of parameters." The fitted values will not generally be exactly equal to the observed data values, and a measure of how discrepant they are provides a good measure of how well the model fits the data. Of the many measures of discrepancy, the one of primary concern for **SURPH** modeling is formed from the logarithm of the ratio of likelihoods, and is known ***as the deviance.***

When observations are made on individual subjects, models fit to a set of N observations can have up to N parameters. The simplest model, ***usually*** known ***as the null model,*** has only one parameter; it fits a common $\hat{\mu}$ to all observations. The null model is also sometimes referred to as ***the grand mean model.*** At **the** other extreme, **the** *full* or ***saturated model*** **has** N parameters, one parameter per observation, and generates $\hat{\mu}$'s that match the data exactly. For most data sets, the grand mean model is too simple, while the saturated model is uninformative because it does not summarize the data but merely repeats them in full. Most practical models will be intermediate between the null and full models, containing p parameters $(1 < p < N)$. ***The*** basis for analysis of deviance is that the null model is nested in any intermediate model, and any intermediate model is nested in the full model.

The log-likelihood for the full model is the maximum attainable by any model for the data, and that for the null model is the minimum for any model. The two extremes provide a way to assess the discrepancy of an intermediate model. The deviance of a model is defined as twice the difference between the maximum log-likelihood attainable (full model) and that attained by the model under investigation. Symbolically,

$$D_M = 2 \{ lnL(\hat{\underset{\sim}{\theta}}_F) - lnL(\hat{\underset{\sim}{\theta}}_M) \} \tag{2.30}$$

where D_M is the deviance for Model M (the model under investigation), $\hat{\underset{\sim}{\theta}}_i$ **are** the **MLEs** under model i and $lnL(\hat{\underset{\sim}{\theta}}_i)$ is the maximum log-likelihood value under model i ($i = F$ or M).

The deviance of ***the*** null model (D_0) is known ***as the total discrepancy*** **(McCullagh** and Nelder 1983). The fit of an intermediate model can be assessed by examining how much of the total discrepancy it “explains” (how much smaller is its deviance than the total discrepancy), taking into account the number of parameters required to achieve the reduction in deviance. For normal-theory linear models, the deviance for a particular model has an exact χ^2 distribution when the model is true. The deviance has an asymptotic χ^2 distribution for distributions other than normal, though **McCullagh** and Nelder (1983) warn that the χ^2 approximation for the deviance may be appropriate only for very large sample sizes. In the context of survival-effects models for tag-release data, Smith (1991) shows that an F-statistic derived from analysis of deviance was nominally distributed in a wide variety of scenarios with moderate release sizes.

Deviance has the attractive property of additivity for nested sets of models. For example, consider the nested set of models $(M_0, M_1, \ldots, M_F)$, where M_0 is the null model, M_F is the full model, and N_m is the total number of models under investigation. Model M_0 is nested in (a special case of) M_1, which is nested in M_2, and so on. Let ν_i be the number of parameters in M_i. The difference between the deviances of M_{i-1} and M_i is equal to the likelihood ratio statistic for testing M_{i-1} against M_i:

$$D_{i-1} - D_i = -2\{lnL(\hat{\underset{\sim}{\theta}}_{i-1}) - lnL(\hat{\underset{\sim}{\theta}}_i)\} .$$

Accordingly, the difference in deviances is denoted $-2\Lambda_i$. The total discrepancy:

$$D_0 = -2\ \{lnL(\hat{\underset{\sim}{\theta}}_F) - lnL(\hat{\underset{\sim}{\theta}}_0)\},$$

can also be **written** as the sum of the N_m likelihood ratio statistics:

$$D_0 = \sum_{i=1}^{N_m} -2\Lambda_i .$$

This decomposition of the total discrepancy suggests the construction of a table similar to the familiar normal theory method of analysis of variance **(ANOVA)** (Table 2.3). ANODEV decomposes discrepancies in the same way that **ANOVA** decomposes sums of squares, leading to an ANODEV table that is analogous to the **ANOVA** table. Extending the analogy to regression analyses, note that **ANOVA** tables are often used to display regression results. An auxiliary statistic is the "coefficient of determination," or R^2. ***An analog*** of R^2 based on deviance is given by the percentage of the total discrepancy that is explained by the model under investigation.

Table 2.3 Basic analysis of deviance table partitioning total deviance into model and residual deviance

Source	d.f.	Change in Deviance	DEV/d.f.	F
Total	$\nu_F - 1$	$D_o = 2\{lnL(\hat{\theta}_F) - L(\hat{\theta}_0)\}$		
Fitted Model	$\nu_M - 1$	$D^*_M = 2\{lnL(\hat{\theta}_M) - lnL(\hat{\theta}_0)\}$	$MD_M = D_M/(\nu_M - 1)$	$F_{\nu_M - 1, \nu_F - \nu_M} = \frac{MD_M}{MD_E}$
Residual	$\nu_F - \nu_M$	$D_E = 2\{lnL(\hat{\theta}_F) - lnL(\hat{\theta}_M)\}$	$MD_E = D_E/(\nu_F - \nu_M)$	

* D_M= can be further decomposed into a sum of single degree of freedom deviances

Symbolically, this is written as

$$R_M^2 = \frac{D_0 - D_M}{D_0} \tag{2.31}$$

Alternative models may be informally compared, based on the proportion of total discrepancy explained by each model.

2.5 Numerical Optimization

To use the maximum likelihood estimation, methods are needed to find the maximum point of the likelihood (or log-likelihood) function. The method illustrated in Section *2.3* involves finding the roots of the system of equations obtained by differentiating the log-likelihood with respect to each parameter. This might be called the "analytical" optimization method. It leads to general algebraic expressions for the MLEs when the roots can be found analytically. Section 2.3 shows how the covariance matrix for the MLEs found by this method can be derived from the matrix of second partial derivatives of the log-likelihood (Information matrix).

There are many examples, however, of likelihood functions that lead to likelihood equations that cannot be solved analytically, or where it is prohibitively difficult to derive the Information matrix analytically. The solution in these situations is to use numerical methods to maximize the function. Numerical methods may be desirable even when analytical solutions are known to exist, but when the analytical solutions are too difficult to derive.

The best known numerical method is the Newton-Raphson (N-R) method. Other methods are variations on or approximations to the N-R method, and are sometimes referred to as quasi-Newton (q-N) methods. These methods are iterative: starting with an initial guess of the location of the maximum (the root of the likelihood equations), the methods use information obtained from evaluations of the log-likelihood function and possibly its derivatives to update the guess. The guess is improved iteratively until the series of parameter estimates converges. The N-R method and q-N methods are described below. The advantages and drawbacks of the various numerical methods are discussed. Program SURPH relies exclusively on q-N methods because of the complexity of the likelihood models.

2.5.1 Newton-Raphson method

While the quasi-Newton methods are designed expressly for the problem of optimization, the Newton-Raphson method was devised for the more general problem of finding roots of a system of nonlinear equations. Thus, the N-R method is applied to the likelihood equations rather than the log-likelihood itself.

Suppose we wish to find the single parameter value θ_0 that maximizes the likelihood function $L(\theta)$. This is equivalent to finding the root of the likelihood equation $\frac{d}{d\theta}\log L(\theta)$. Let $g(\theta) = \frac{d}{d\theta}\log L(\theta)$. If we start with a trial solution $\theta_{(1)}$ in the neighborhood of θ_0, the Taylor approximation gives:

$$0 = g(\theta_0) \approx g(\theta_{(1)}) + (\theta_0 - \theta_{(1)})\, g'(\theta_{(1)})$$

where g' (θ) is the first derivative of g (θ) (i.e., the second derivative of $\log L(\theta)$) with respect to 8, evaluated at $\theta_{(1)}$. Rearranging this equation gives a new approximation to θ_0 :

$$\theta_0 \approx \theta_{(1)} - \frac{g(\theta_{(1)})}{g'(\theta_{(1)})} .$$

Thus, on each step of the iteration, a new trial solution is obtained:

$$\theta_{(i+1)} = \theta_{(i)} - \frac{g(\theta_{(i)})}{g'(\theta_{(i)})} .$$

Iteration continues until $|\theta_{(i+1)} - \theta_{(i)}| < \varepsilon$ for some predetermined ε.

Suppose now that the parameter, $\underset{\sim}{\theta}$, is a vector of length p and we wish to find the maximum point $\underset{\sim}{\theta}_0$ that maximizes $L(\underset{\sim}{\theta})$. *The likelihood* equations are $\underset{\sim}{g}(\underset{\sim}{\theta})$, where

$g_i(\underset{\sim}{\theta}) = \frac{\partial}{\partial\theta_i}\log L(\underset{\sim}{\theta})$. *The* function g $(\underset{\sim}{\theta})$ is sometimes called the "score vector." Starting with an initial trial solution vector, the multivariate version of the N-R method is

$$\underset{\sim}{\theta}_{(i+1)} = \underset{\sim}{\theta}_{(i)} - H^{-1}_{(i)}\underset{\sim}{g}(\underset{\sim}{\theta}_{(i)}),$$

where $H^{-1}{}_{(i)}$ and g $(\underset{\sim}{\theta}_{(i)})$ are the Information matrix and score vector, respectively, evaluated at $\underset{\sim}{\theta}_{(i)}$. The *jk*th element *of* $H_{(i)}$ is

$$\frac{\partial^2}{\partial\theta_j\partial\theta_k}\log L(\underset{\sim}{\theta})\Big|_{\underset{\sim}{\theta} = \underset{\sim}{\theta}_{(i)}}.$$

Iteration continues until some convergence criterion is satisfied. A typical criterion is $\|\underset{\sim}{\theta}_{(i)} - \underset{\sim}{\theta}_{(i-1)}\| < \varepsilon$ where II •II is the Euclidean norm $\left(\|x\| = \sum x_i^2\right)$.

Implementation of the N-R method in a computer program requires that all the first partial and mixed second derivatives of the log-likelihood be coded. This can be extremely difficult, particularly if the program must be general enough to deal with many different study designs or parameterizations of survival in a multiple-population tag-release model. The derivatives must be calculated on each step of the iterative procedure. On some computers, this computation will be slow, especially if the number of parameters is great. In addition, the derivative matrix $H_{(i)}$ must be inverted on each step. Matrix inversion is a notoriously slow operation, and is susceptible to numerical instabilities if not programmed carefully. A benefit of the method, though, is that at the end of the final iteration, the observed Information matrix is available, which is inverted to estimate the covariance matrix.

2.5.2 Quasi-Newton methods

Quasi-Newton methods are used to avoid the difficulty of deriving and coding the first and second partial derivatives. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method requires only first derivatives (Press et al. 1986). It uses successive evaluations of the log-likelihood and the first derivatives to construct an approximation to the inverse Information matrix. That is, a sequence of matrices $A_{(i)}$ is constructed with, the property

$$\lim_{i \to \infty} A_{(i)} = [I(\hat{\underset{\sim}{\theta}})]^{-1} .$$

Because convergence (it is to be hoped) occurs within a finite number of iterations, the value of A from the final iteration $(A_{(N)})$ is an estimate of the covariance matrix. The trial solution vector is updated as

$$\underset{\sim}{\theta}_{(i+1)} = \underset{\sim}{\theta}_{(i)} - A_{(i)} \underset{\sim}{g}(\underset{\sim}{\theta}_{(i)}) .$$

Details of how $A_{(i)}$ is constructed and updated may be found in Press et al. (1986, pp. 307-310).

A second quasi-Newton method, due to Fletcher (1970) and known as "FLETCH," uses no explicit evaluations of first or second derivatives. Instead, only the log-likelihood function is used, and difference equations are used to construct a series of approximations to the score vector and the Information matrix $I(\hat{\underset{\sim}{\theta}})$.

Code is available from outside sources for the quasi-Newton methods (Press et al. 1986, for BFGS; Vaessen 1984 for FLETCH). Of course, the user must provide code to evaluate the log-likelihood for both methods, and the first derivatives for BFGS. The FLETCH code appears to be more self-contained and readily adaptable. To build up reasonable approximations to the Information matrix, both BFGS and FLETCH require many more function evaluations than does N-R. The advantage is that coding of derivatives is avoided. Both programs provide, as by-products of the final iteration, the maximized function value for use in LRTs and a data structure that contains the estimated covariance matrix or the data that may be used to construct it.

Clearly, the main desideratum in selecting a numerical optimization method for a particular problem is the trade-off between coding derivatives and the number of function and derivative evaluations required. Although each iteration of the N-R algorithm will take more CPU time, the q-N methods will usually require so many more iterations and function evaluations per iteration that the total CPU time is greater for q-N. However, since CPUs are becoming exceedingly fast, this need not be a great concern. Because of the difficulty of coding second derivatives for the

SURPH models, particularly models for release-recapture data, the N-R method is not practical. Among the q-N methods, **FLETCH** is used in **Program** SURPH because the available code is much easier to use. One consequence of using **FLETCH** is that occasionally, **SURPH** will not find the **MLEs** because the method fails to converge. Invariably, failure to converge indicates that the data are too sparse to estimate all the parameters of the model being fit.

Chapter 3

Installing Program SURPH and Creating Data Files

3.1 System Requirements

To run SURPH under the UNIX operating system, the **following** equipment and support are necessary:

1. A SUN workstation with at least 7 MB of free memory. The amount of free memory above 7 MB will determine the speed of the program execution.
2. The UNIX operating system.
3. Some version of an X-windowing system. SURPH has been run using **XView** and several other X-window managers (e.g., uwm, twm). It should run under any window manager that uses X.
4. A color monitor. Although color is not necessary, it is helpful in viewing certain graphics.

To run SURPH under the Microsoft Windows@ operating system, the following equipment and support are necessary:

1. An IBM-compatible personal computer with at least 7 **MB** of free memory.
2. The Microsoft Windows@ operating system.
3. A color monitor. Although color is not necessary, it is helpful in viewing certain graphics.

3.2 Installing SURPH

Step One: Accessing the executable and other files.

Use anonymous ftp or Mosaic to log in to the host site and download the necessary files. The **address** of **the** host site **is ftp://opus.cqs.washington.edu/public/surph**. To get the files via ftp, at the UNIX prompt type

> ftp opus.qs.washington.edu

At the **login** prompt, type "anonymous", then use your **login** id as the password.

To get the files **via Mosaic, type ftp://opus.cqs.washington.edu/public/surph in the** new **URL** pop-up window.

The files are:

1. **README** (ascii mode) Contains these instructions, including any revisions since this manual was printed.

2. **SURPH.tar.Z** (binary mode) Compressed "tape archive." Use uncompress and tar utilities to extract the following files:

    surph (binary mode) Contains the entry program.

    omni (binary mode) Contains the release-recapture program.

    telem (binary mode) Contains the known-fate program.

    ***.info** (ascii mode) These are help files; they include **help_main.info, anodev.info, menu.info, xv.info, desc_stat.info,** and **splots.info.**

3. SURPH.data.tar.Z (binary mode) Compressed "tape archive" containing data **files.**

4. **SURPH.manual.tar.Z** (binary mode) Compressed "tape archive" containing this manual.

Step Two: Setting the PATH environment variable.

Set the PATH environment variable to include the directory that contains downloaded SURPH files.

Step Three: Setting the SURPH environment variable.

Create and set a new environment variable called SURPH. This variable should contain the name of the directory that SURPH is installed in (that is, the same directory that was added to the path variable in the previous step).

Step Four: Setting the **HELPPATH** variable.

Add to (or create) the environment variable **HELPPATH** to include the directory where the *.info files will reside. This directory can be different from the directory that SURPH is installed in, but it doesn't have to be.

After these four steps are performed, SURPH should be operational.

Send questions or bug reports to **surph@cqs.washington.edu**.

3.3 SURPH Data Files

3.3.1 Introduction

This section describes the format required for **SURPH** data files. Whether a study uses release-recapture or known-fate sampling to collect data, the following information must be provided in every SURPH data file:

1. A brief verbal description, or title, for the data set.
2. The number of discrete populations in the data set.
3. The number of sampling intervals (i.e., intervals between sampling events). If there are ***K*** sampling events, then there are K-l sampling intervals.
4. A definition of each group covariate. For each covariate, the data file should include a name, an indication of whether the covariate varies from interval to interval, and the measurements of the variable for each population.
5. The number of animals newly tagged and released in each population at each sampling event except the last (new animals are not tagged on the last sampling events).
6. The number of characters used for individual identification **(ID)** codes. If codes are not listed in the data file, the user should indicate an "ID length" of 0 characters.
7. The name of each individual covariate.
8. For every tagged individual in the study: the identification code, if desired; the capture or survival history; and the measurements for each individual covariate in the same order as the names provided above **(#7)**.

Below are details on how the required information must be arranged in the data **file.** Rules for capture or survival histories are also provided, particularly for animals that are known to have been removed from the population on a particular sampling event.

Figure 3.1 illustrates the **first** section (the "Header") of a SURPH data file for a hypothetical release-recapture study. This simulated data set was used to generate most of the "snapshots" of the program found in Chapter 4. Each piece of information provided to SURPH is **identified** with a "keyword" (Table 3.1). In many cases, several synonymous keywords can be used to identify the same information. Novice users may prefer to use the longer, fully-descriptive keywords, while more experienced users may prefer the shorter, more cryptic ones. Synonymous keywords are grouped together in Table 3.1. In the text that follows, keywords appear in bold type. Examples of such keywords in the **first** few lines of Figure 3.1 are **Data, number-of-populations,** and **group,covariate. Each** data file consists of two sections: the "Header", which begins with the keyword Data and extends to the keyword **data,** where the "Data Block" begins.

3.3.2 Header format

The first line of every data file is the single keyword **Data,** which identifies the file as a data file and distinguishes it from model files produced when fitted models are saved (the **first** line of a model file is the keyword **Models). The** second **line** is the verbal description, or title of the data set. The title appears at the top of most display screens in SURPH, allowing easy identification of the data being analyzed. The title is also saved in model files, so the program can ensure that any model files the user attempts to load are compatible with the **data being** analyzed.

The next two lines define the dimension of the data set. The example data file **(Figure** 3.1) has six populations **("number-of-populations 6")** and four sampling intervals **("number-of-intervals 4').**

Comments can be placed in the data file at any point after the data title. Lines beginning with the character **"#"** are comments, and ignored when SURPH reads data files.

```
Data
Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
# 3 paired releases for total of 6 populations
number_of_populations 6
# 5 sampling events give 4 sampling intervals
number_of_intervals 4

group_covariate name treatment time 0
   0.000000
   1.000000
   0.000000
   1.000000
   0.000000
   1.000000
group_covariate name mid-release time 0
   0.000000
   0.000000
   1.000000
   1.000000
   0.000000
   0.000000
group_covariate name late_release time 0
   0.000000
   0.000000
   0.000000
   0.000000
   1.000000
   1.000000
group_covariate name ATPase time 1
8.3        255     29.1     28.8
9.0        20.42 25.9     33.02
9.3        22.56  34.14  39.28
7.5        20.2    29.0     23.6
8.5        21.4    34.0     25.3
9.0        21.3    23.9     23.6

individual~covariate  length
individual~covariate  weight
individual~covariate  condix
number_tagged
1000 250 0 0
1000 250
1000 250 8 8
1000 250    0
1000 250 8 0
1000 250 0 0
length_of_ids 10
data
7F7E644A5A       10000       141 31.4 1.12
7F7F0F554E       10000       144 33.6 1.12
# 7F7F0F5C7D transported by barge from Lower Granite Dam
7F7F0F5C7D      12
7F7F103974          10100    156 146 39.9 34.9 1.04 1.11
7F7F112C28
7F7F112F77      11000 102    139 124 29.5 21.3 1.09 1.12
7F7F15131F      01010        132 24.6 1.08
```

Figure 3.1 Header and beginning of data block for hypothetical release-recapture data file.

Table 3.1 Data file keywords.

Keyword	**Purpose**
Data	First word on **first** line of every file. **Identifies** file as a data file.
number-of-populations **npop**	Number that follows the keyword is the number of populations.
number_of_intervals nint number-of-periods **nper**	Number that follows the keyword is the number of sampling **intervals**.
group-covariate **gcov**	First word on line that **describes a** group covariate.
name	On gcov line; character string that follows the keyword **name** is the name for the group-based covariate.
time	On gcov line; number that follows the keyword **time** is 0 if covariate is time-independent and 1 if time-dependent.
individual-covariate **icov**	Character string that follows the keyword is the name of an individual-based covariate.
number-tagged **ntag**	Block of numbers that follows the keyword is the number of newly-marked animals released in each population at each sampling event.
full_hist	Must appear on a separate line before the keyword data if capture histories for removed animals are extended with O's **after the** removal event. **The** keyword **full,hist** is absent if capture histories of removed animals are truncated after the removal.
length-of&is **idlen**	Number that follows the keyword is the length (number of characters) of the code used to identify individual animals.
data	Identifies beginning of Data Block. Data that follow are ID codes, capture or survival histories, and covariate measurements for individual animals.
enddata	Identifies end of Data Block

In the example data set, the six populations represent three paired treatment/control releases. The next section of the data file defines group **covariates** that assign each population to the appropriate treatment level and pair, allowing modeling of the effects of treatment and blocking factors. The treatment and release-pair covariates are time-invariant indicator **variables.** For example, the line **"group_covariate name** treatment time 0" defines a time-invariant group covariate named "treatment." The next six lines give the covariate value for each of the six populations. Because the covariate is time-invariant, there is one value on each line. Populations **2, 4,** and 6 **are** the treatment populations. The release-pairs are defined similarly, with group covariates named "mid-release" and "late-release." Individuals from each population were sampled at each sampling event, and the sample average **ATPase activity used as** a group covariate to characterize the entire population. The fourth group covariate line **("group-covariate name ATPase time** 1") defines the **ATPase** variable, Because the covariate changes between sampling intervals **(time 1),** the next six lines give the covariate values for each interval for each population.

For each individual covariate, the **data** file contains one **line (individual,covariate** ***covariate_name*)** to define the covariate name. In the example data set, measurements of length, weight, and condition index (e.g., "condix") are defined. The covariate values for each individual are provided below in the Data Block.

The numbers of animals marked and released at each sampling event in each population are in a block following the keyword **number-tagged. In the** example, 1000 animals were released on the first event in each population, and 250 were released on the second event. No newly marked animals were released on the third or fourth events.

The optional keyword **full_hist** is used if capture or survival histories (see Sections 3.3 and **3.4)** for censored animals are extended with O's after the removal event. If the capture or survival histories are truncated after the removal event, the keyword **full_hist is** omitted.

The keyword **length-of-ids** must appear at some point in every SURPH data file. Following the keyword is the length (number of characters) of the ID code used for each individual. In the example, **ID codes are** 10 characters long **(length_of_ids** 10). If no ID codes are listed in the **data** file, **the** "length" is 0 **(length-of-ids** 0).

The keyword **data** indicates the end of the Header and the beginning of the Data Block.

3.3.3 Data Block format

The keyword **data** signals the beginning of the Data Block. The Data Block contains one line for each animal marked during the study. In the example (Figure **3.1),** the total number of animals from all six populations is 7500; hence, there are 7500 lines in the Data Block. The record for each individual includes the **ID code, if used (see length_of_ids** above); the capture or survival history, with digits separated by spaces; and the values of each individual covariate measured on the individual, listed in the order in which the covariate names were defined in the Header. For example, the individual with PIT-tag code **7F7F103974** (Figure 3.1) was 146 millimeters long, weighed 34.9 grams, and had a condition index of 1.1. See Sections 3.3 and 3.4 for explanations of the capture and survival histories.

Individual records within the Data Block are ordered by population. That is, all records for Population 1 appear before records for Populations **2, 3,** and so on. Thus, the example file would include first the records for the 1250 animals tagged in Population 1, then the 1250 for Population 2, and so on. Within the sections for each population, the records for the individual animals can be listed in any order; records need not be listed in the order that the individuals were marked.

The last line of every data file is the keyword **enddata,** signaling the end of the Data Block and the end of the entire file. SURPH provides notification of **an** error **if enddata** is encountered "too soon" or "too late" (i.e., the number of records in the Data Block is not equal to the total number of marked animals declared in the **ntag** block of the Header).

3.4 Rules for Survival Histories (Known-Fate Studies)

A survival history is the record of the sampling events at which an individual was detected alive. Survival histories for individuals that were not known to be removed consist of a series of

K digits, where *K* is the number of sampling events. If an individual is detected alive, a "1" is recorded in the survival history for the particular interval. If an animal is detected but is found dead, a **"0"** is recorded. Individuals not released at the **first** event, but captured and marked at the second event have survival histories beginning with the digits "0 1." For example, an individual marked at the second event, then redetected at the third and fourth events, but not at the fifth event, would have a survival history of "0 1 1 10." Survival histories for animals newly marked at the third event begin with "0 0 1," and so on.

In known-fate studies, each marked individual is queried at every event, and it is determined whether the animal has survived or died since the previous event. Thus, survival histories in known-fate studies cannot have "O's" between two "l's"; the first **"0"** following the release occasion indicates mortality. Thus, an individual in a known-fate study could have **a survival** history of **"0** 0 1 **1 0",** which indicates release on the third occasion, detection on the fourth, and then death. However, an animal in a known-fate study could not have a survival history of "0 10 0 1" for example, because the zero on the fourth event indicates that the animal died; it could not then be redetected on later sampling events.

There are two distinct types of known removals in known-fate studies. For example, in a study using radio-tags, animals are removed from the sample **if**: (1) they are captured alive and removed from the study site (e.g., handling mortality or for biopsy); or (2) their radio transmitter fails and their survival status cannot be determined. Survival histories for removals of the first type are recorded the same as removals for release-recapture, with a code of "2" for the removal event. For example, in a study including five sampling events, if an animal survives to the third event and is removed on that event, the truncated survival history is "1 12" **("1** 12 0 0" if extended: **full,hist** keyword included). For removals of the second type (i.e., radio transmitter failure), a code of "3" is recorded on the first occasion on which the individual could not be located. In the five-sampling event study above, an animal that was known to survive until the third event, but could not be located on the fourth event or thereafter, would have a survival history of "1 1 13" **("1** 1 13 **0,"** if extended). With regard to information provided for survival estimation, survival histories "1 12" and "1 1 13" are equivalent; both animals provide information for the first two sampling intervals but none for any subsequent interval.

3.5 Rules for Capture Histories (Release-Recapture Studies)

A capture history is the record of the sampling events on which the individual was either detected alive or not detected. Capture histories for individuals that were not censored consist of a series of K digits, where K is the number of sampling events. If an individual is detected, a **"1"** is recorded in the capture history for that particular event. If not detected, a "0" is recorded. In the example file (Figure **3.1)**, the individual coded **7F7F103974** was marked and released at the first event, redetected at the third event, but not detected on events **2, 4,** or 3. Thus, the capture history is "1 0 10 0." Individuals not released at the first event, but captured and marked at the second event have capture histories beginning with the digits **"0** 1". For example, individual **7F7F15** 131F (Figure 3.1) was marked on the second event and then redetected on the fourth event, for a capture history of **"0** 10 lo". Capture histories for animals newly marked on the third event begin with the digits "0 0 **1",** and so on.

In release-recapture studies, a failure to detect a marked individual may be due to mortality or to imperfect sampling of surviving individuals. Thus, capture histories typically consist of a series of interspersed "1 's" and "O's" (see example data set, Figure 3.1).

There are a number of ways that marked animals may become known removals from a release-recapture study. For example, if physical trapping is used, animals may be killed by trapping or handling; or live animals may be purposely removed for biopsy. In the automatic data collection system for PIT-tags on the Snake River, the system fails to divert a portion of the tagged fish back into the river. These fish are transported by barge, rather than returned to the river. For known removals, the code "2" is used to denote the event when the animal is removed. The code "2" indicates that the individual survived the interval but was removed from the study before contributing survival information for subsequent intervals. For example, individual **7F7F112C28** (Figure 3.1) was removed after being detected alive at the third event, for a truncated capture history of "10 2".

Often, SURPH users will record capture data in a spreadsheet, with one column for each sampling event and one row for each marked animal. The Data Block for the data file can then be generated by saving the spreadsheet in a text format. For data files produced in this way, it can be convenient to present capture histories for removed individuals with “extended O’s” after the removal event, so there is a numeric code for&h event. For example, individual **7F7F112C28** (Figure 3.1) would have an extended capture history of “10 2 0 0”. If capture histories in the Data Block **are** extended, **the** keyword **full,hist** must appear in the **Header. If full-hist** does not appear in the Header, **SURPH** assumes that capture histories are truncated after **removal** events.

Chapter 4

Using Interactive Program SURPH

Program SURPH provides an interactive windows environment for summarizing and analyzing known-fate and release-recapture data. This chapter provides instructions on how to use the interactive program, windows and cursor/mouse (Section 4.1). Furthermore, this chapter provides an illustrated index of the program options, directives and data displays. Section 4.2 provides guidance for the use of Program SURPH in the analysis of release-recapture. This section also describes the use of **SURPH** in the analysis of known-fate data where the operations are the same as for release-recapture data. Section 4.3 provides additional descriptions of the program options unique to the analysis of known-fate data.

This chapter was written primarily for those using the UNIX operating system. The Microsoft **Windows**® version of the program for the PC looks and operates in a very similar fashion. Differences between the two versions of SURPH (i.e., **UNIX** versus Windows*) are noted in a separate section at the end of this chapter.

4.1 Fundamentals in Windows operations

We have included an introductory section that outlines how to use a mouse, how to re-size windows, how to access pull-down and pull-right menus, and how to store text pane information. Also included within this section are descriptions of some conventions and definitions used throughout Chapter 4.

4.2 Analysis of release-recapture data

Program options available for the analysis of release-recapture data are summarized in Table 4.1. The order of the listing is based on a typical sequence of data analysis steps. Specifically, the ordering includes file input/output, descriptive statistics of the data set, modeling of capture and survival probabilities, and model diagnostics. Specific program options are cross-listed in the index of this manual.

4.3 Analysis of known-fate data

The majority of program options available for the analysis of known-fate data are identical to program options for release-recapture data. Only options unique to known-fate data are included in Section 4.3 and summarized in Table 4.2. For all other program options, consult Section 4.2 and Table 4.1. The order of the listing is based on a typical sequence of data analysis steps. **Specifically,** the ordering includes descriptive statistics of the data set, modeling survival probabilities and model diagnostics. Specific program options are cross-listed in the index of this manual.

Table 4.1: Listing of SURPH options for analysis of release-recapture data and associated page number where operational description can be found.

Table 4.1 (cont'): Listing of SURPH options for analysis of release-recapture data.

Table 4.1 (cont'): Listing of SURPH options for analysis of release-recapture data.

4.3 Analysis of known fate data

The majority of program options available for the analysis of known fate data are identical to program options for release-recapture data. Only options unique to known-fate data are included in Section 4.3 and summarized in Table 4.2. For all other program options, consult Section 4.2 and Table 4.1. The order of the listing is based on a typical sequence of data analysis steps. Specifically, the ordering includes descriptive statistics of the data set, modeling survival probabilities and model diagnostics. Specific program options are cross-listed in the index of this manual.

Table 4.2: Listing of unique SURPH options for analysis of known fate data and associated page number where operational description can be found.

Program Options	Page
SURPH Base Window: Known-Fate	4.181
Data Pull-Down Menu	4.185
Data: Binomial Estimates	4.187
Data: Sampling Summary	4.189
Survival Modeling: Pull-Down Menu	4.191
Models Pull-Down Menu	4.193
Models: Model Graphics: Binomial Estimates	4.195

4.4 SURPH-PC

The majority of program options available for the UNIX version of SURPH are duplicated in SURPH-PC. Only options unique to SURPH-PC are included in Section 4.4. For all other program options, consult Sections 4.2-4.3 and Tables 4.1-4.2.

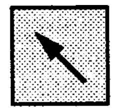
Table 4.3: Listing of unique SURPH options for SURPH-PC and associated page number where operational description can be found.

Program Options	Page
SURPH-PC Introduction	4.197
Differences in SURPH-PC	4.197

Table 4.3 (cont'):

All functions within SURPH are accessed via the mouse. There are only a few times that you must type information while in SURPH. You, therefore, must be thoroughly familiar with mouse operations before any-data analysis is performed using SURPH.

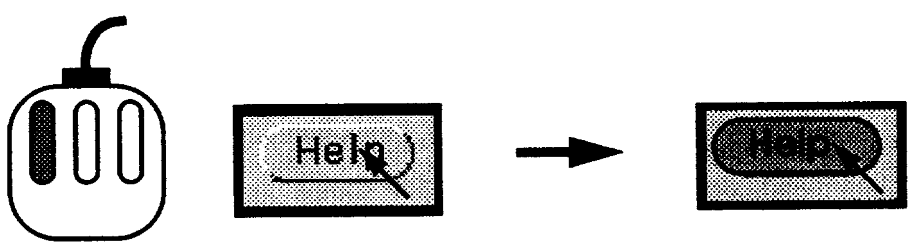
When using the mouse, a floating black arrow (the pointer) indicates your position on the screen or within the window.



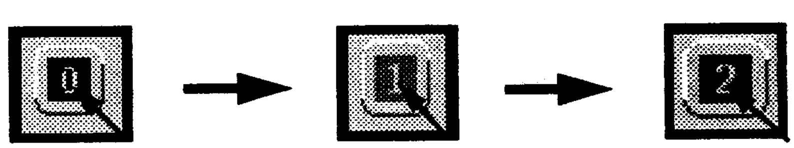
The mouse is used to move the pointer around the screen, to select buttons in SURPH, to access menus, to dismiss pop-up windows, to resize text panes, and to place the cursor in text fields.

In the instructions that follow, we use several terms to describe the functions of the three-button mouse:

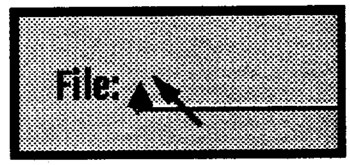
"Left-click": When the pointer covers an object such as a button or a text field, depress and release the left mouse button to select this object. In SURPH, when a button is selected, the button gets darker or the number within the button changes. In the **first** example, left-clicking on **the Help Button** causes the button to darken, and accesses **the** "Help" function.



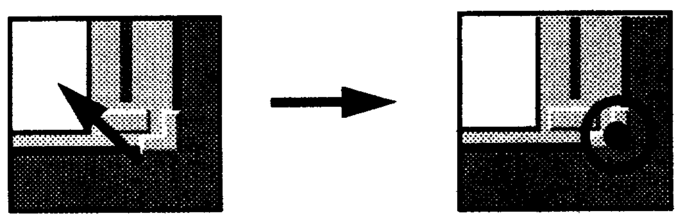
In the second example, left-clicking on a button pad increases the value of the button by one. This type of button is used for model **parameterization.**



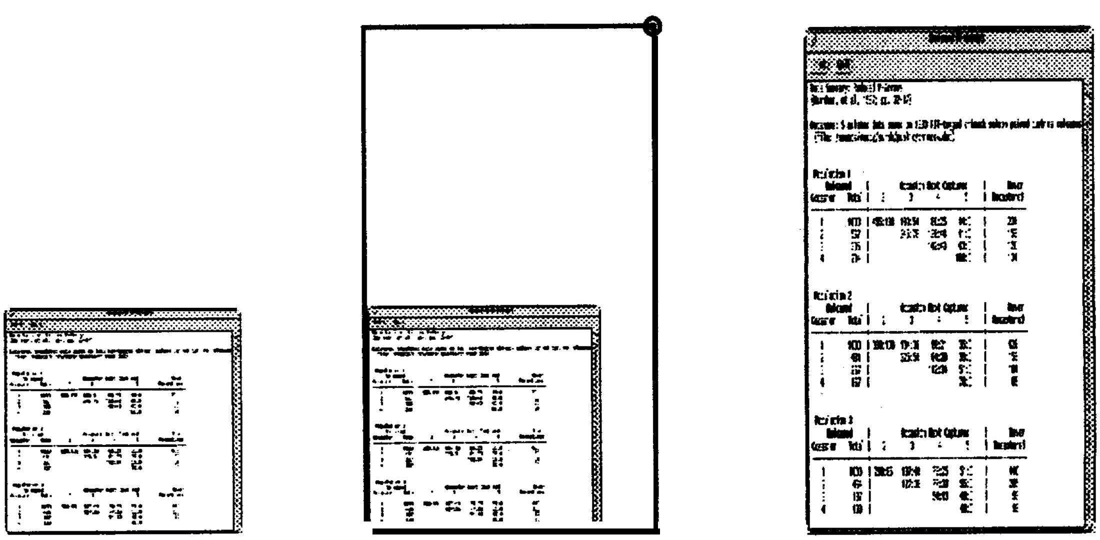
To type in a text field, use **the** mouse to point at the text field, then left-click. A triangular cursor appears within **the** text field at the pointer position. In the following example, you could now enter a file name in the text field.



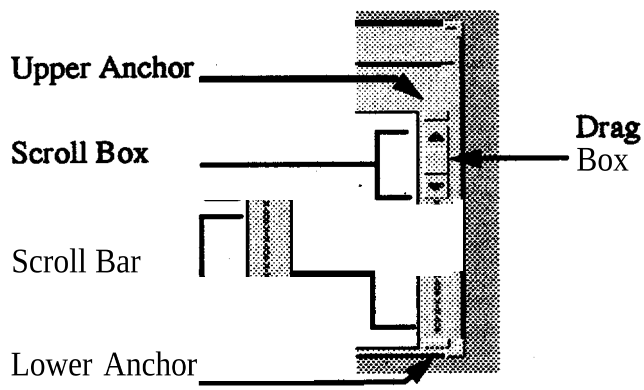
“Left-hold”: Left-hold is used to change the size of windows that allow re-sizing, and to move **the** scroll bar on text windows. When the pointer covers the comer of a text window, it changes from an arrow to a small circle. To resize a text window, move the pointer to one corner, until the **arrow** becomes a circle, as shown here.



Depress the left mouse key and hold down. When you move the mouse while holding the left button, the corner of the window will move **with** you. This allows you to alter the dimensions of the window. Release the mouse button when the window is **the** desired size. In the example, we enlarge an M-Array text window.



All text windows have a scroll bar that extends vertically along the right-hand side of the border, a scroll box that consists of an up-arrow, a down-arrow, and a drag-box; and upper and lower anchors. These scrolling devices are used to move through the text pane from one screen display to the next. The scrolling devices allow you to move through the text pane in four different ways.



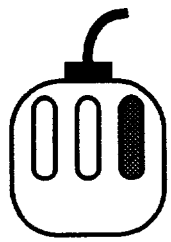
1. To scroll through a document one screen (page) at a time, left-click while the arrow is on the scroll bar. The scroll bar allows you to move either up or down within the document. To move down (up), make certain the arrow is below (above) the scroll box. To move through multiple pages, use left-hold instead of left-click.

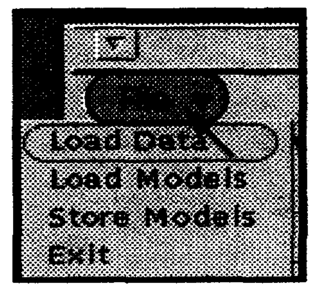
2. To scroll through a document one line at a time, left-click while the arrow is one the scroll box. The scroll indicator allows you to move either up or down within the document. To move down (up), make certain the arrow is on the scroll-box arrow pointing down (up). To move through multiple lines, use left-hold instead of left-click

3. To go to the beginning (end) of a document displayed in a text window, left-click when the arrow is on the upper (lower) anchor.

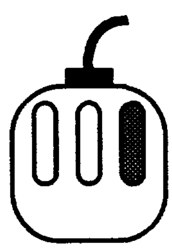
4. An alternative way to move multiple lines or pages is to use the drag box. Place the mouse arrow on the drag box and left-hold while moving the drag box up or down the scroll bar.

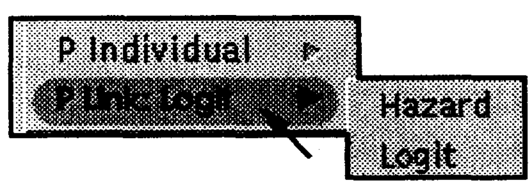
"Right-click": When the pointer covers a button or menu choice that has a small triangle that points downward, depress and release the right mouse button to access a pull-down menu. When a button is selected, the button gets darker. **In the** example, when **you right-click on the File Button,** a pull-down menu appears that allows you to select one of four options. You would then left-click on the option you wish to use.



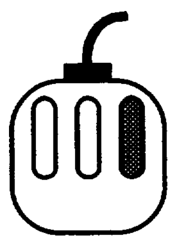


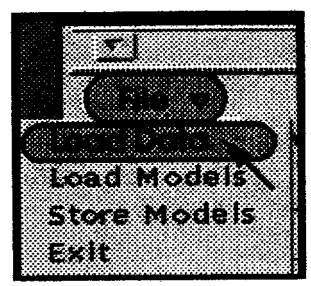
When the pointer covers a button or menu choice that has a small triangle that points to the right, depress the right mouse button to access a **pull-right** menu. In this example, when you right-click on **P Link: Logit,** a pull-right menu appears that **allows you to** select either **the Hazard or Logit link**. **The P Individual** is shown to illustrate the difference in shading between selected and non-selected choices.





"Right-hold": Right-hold is used to perform **the** same mouse operations as right-click. However, instead of depressing the right mouse button and releasing, depress the right mouse button and hold. While holding **the** right button down, move the mouse until the option you are interested in is selected (i.e. the option is darkened), then release the right mouse button. In the example, the right-hold **was used** to access the **File Button** pull-down menu. While holding the right mouse button, the pointer was moved to **Load Data** and released. This action implemented the **Load Data** function.





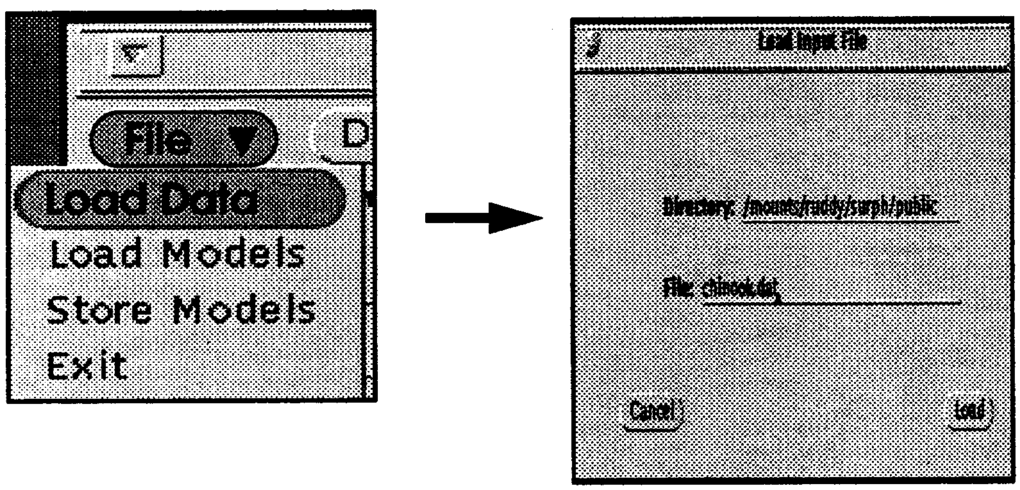
The middle button is not used for anything in **SURPH.**[1]

1. **For further instructions on how to use the mouse, consult the section on using the mouse in your User's Guide.**

THIS PAGE INTENTIONALLY LEFT BLANK

Conventions

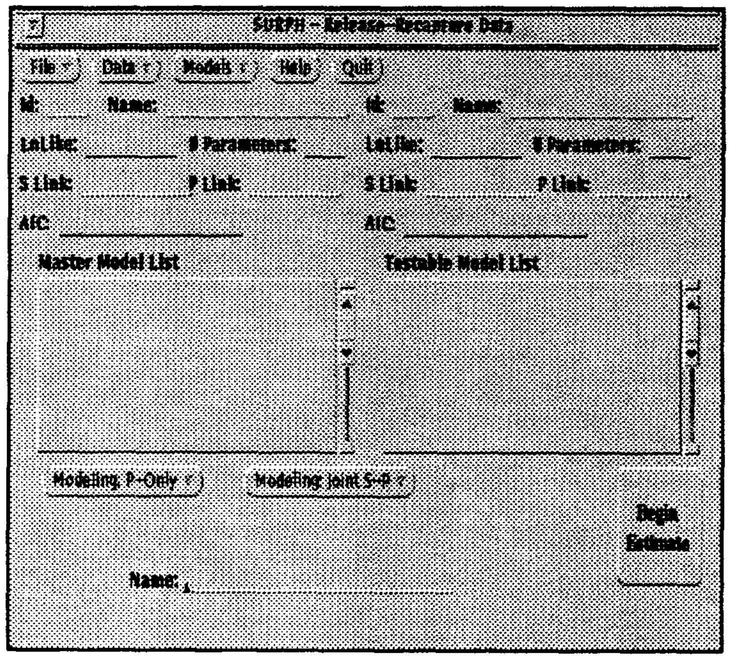
All commands in SURPH are accessed using **the** mouse keypad. Throughout this chapter, command paths will be included along with captured screen graphics to facilitate learning SURPH. The “command paths” are the listings of the necessary selections you must make to obtain the graphical display shown on that page. For example, **to** obtain the Load Data Window, you must right-click on **File,** then left-click on **Load Data. This** command path **is** illustrated **as File -> Load** Data. The result of this command path is **the** appearance of the Load Data Window.



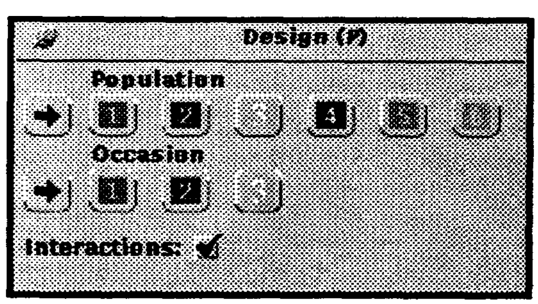
Definitions

-

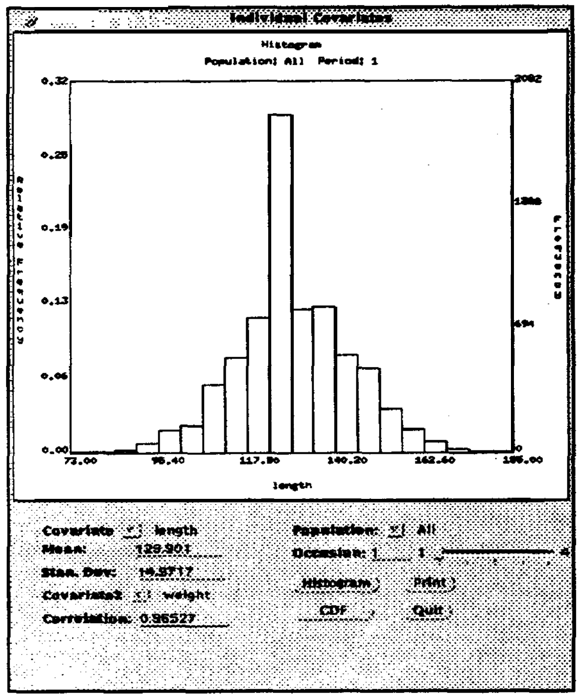
Button - A Button is a device that allows you to implement functions.-Examples of Buttons on the SURPH Base Window are **File, Data, Models, Help, Quit, Modeling: P-Only, Modeling: Joint S-P,** and **Begin Estimate.**



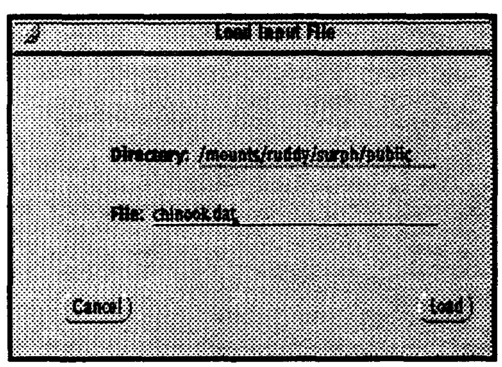
Button Pad - A Button Pad is a device that allows you to change the model parameterization. Button Pads are defined for all Capture (P), all Survival (S), and all Product (SP) design factors, group covariates and individual covariates. Left-click on a button within the button pad to change the model parameterization. For further Button Pad illustrations, see Quick Button Pad Instructions



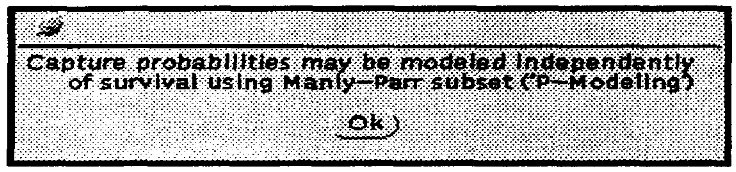
Graphics Window - A Graphics **Window** displays information graphically. An example of a Graphics Window is the **Pop-Up** that appears after you selects **Data -> Individual Covariates.**



Input/Output Window - An Input/Output Window in SURPH allows you to load data and to load or store models. An example of an Input/Output **Window** is the Load Data Window under **File -> Load Data.**

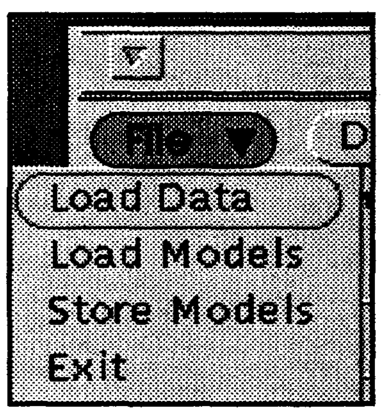


Message Window - A Message **Window** in SURPH is used to query or enlighten you. An example of a Message Window **is the** Pop-Up that appears after you initially choose to model capture probabilities **under the Modeling: Joint S-P Button.**

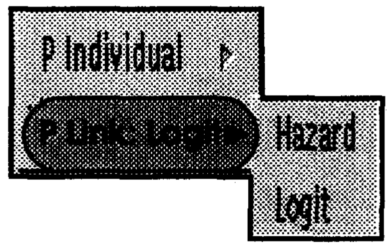


Pop-Up Window - A Pop-Up Window in SURPH is any window that appears following a command. Examples of Pop-Up Windows are Message Windows that query or enlighten you, Text Windows that present text information, Graphics Windows that present graphical information, and Input/Output Windows that allow you to store or load files. Pop-Ups are usually used for temporary tasks or displays. All Pop-Up Windows have a push-pin in the upper left-hand corner. To dismiss a Pop-Up window, left-click on the push-pin.

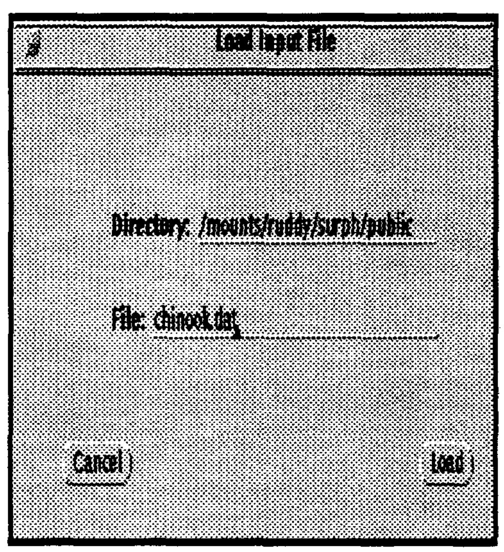
Pull-Down Menu - A Pull-Down Menu gives you additional options under a Button. Pull-Down Menus are accessed by right-clicking on a Button with a triangle that points downward. Examples of Buttons with Pull-Down Menus are **File, Data, Models, Modeling: P-Only,** and **Modeling: Joint S-P.**



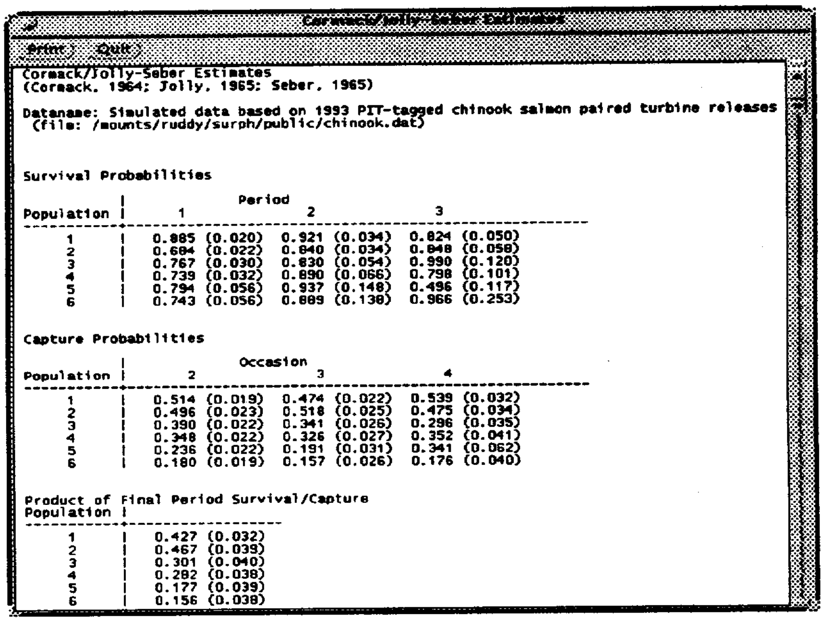
Pull-Right Menu- A Pull-Right Menu gives you additional options under a **Pull-**Down Menu choice. Pull-Right Menus are accessed by right-clicking on a Pull-Down Menu option with a triangle that points to the right. An example of a Pull-Down Menu with a Pull-Right Menu **is Modeling: P-Only -> P Link: Logit**.



Text Field - A Text Field is a region on a Pop-Up Window that allows you to enter text, usually a directory path or a file name. Text fields are used in SURPH to name models, to name new variables, to load data, and to load and store models. In the example, both the underlined area following **Directory:** and **File: are** text fields that can be edited by you. In this particular example, you has instructed the program to look for the data file named **"chinook.dat"** in the directory **"/mounts/ruddy/surph/public"**.



Text Window - A Text Window displays text information. **In** SURPH, text windows are used to display various **summary** statistics and model estimates for **the** current data set. An example of a text window is the window that displays **Cormack/Jolly-**Seber Estimates **(Data -> CJS Estimates). This** window displays estimated capture and survival probabilities based on the model **parameterization** of Cormack **(1964),** Jolly (1965) and Seber (1965).



Print Quit

Cormack/Jolly-Seber Estimates
(Cormack, 1964; Jolly, 1965; Seber, 1965)

Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
(file: /mounts/ruddy/surph/public/chinook.dat)

Survival Probabilities

	Period		
Population	1	2	3
1	0.885 (0.020)	0.921 (0.034)	0.824 (0.050)
2	0.684 (0.022)	0.840 (0.034)	0.848 (0.058)
3	0.767 (0.030)	0.830 (0.054)	0.990 (0.120)
4	0.739 (0.032)	0.890 (0.066)	0.798 (0.101)
5	0.794 (0.056)	0.937 (0.148)	0.496 (0.117)
6	0.743 (0.056)	0.889 (0.138)	0.966 (0.253)

Capture Probabilities

	Occasion		
Population	2	3	4
1	0.514 (0.019)	0.474 (0.022)	0.539 (0.032)
2	0.496 (0.023)	0.518 (0.025)	0.475 (0.034)
3	0.390 (0.022)	0.341 (0.026)	0.296 (0.035)
4	0.348 (0.022)	0.326 (0.027)	0.352 (0.041)
5	0.236 (0.022)	0.191 (0.031)	0.341 (0.062)
6	0.180 (0.019)	0.157 (0.026)	0.176 (0.040)

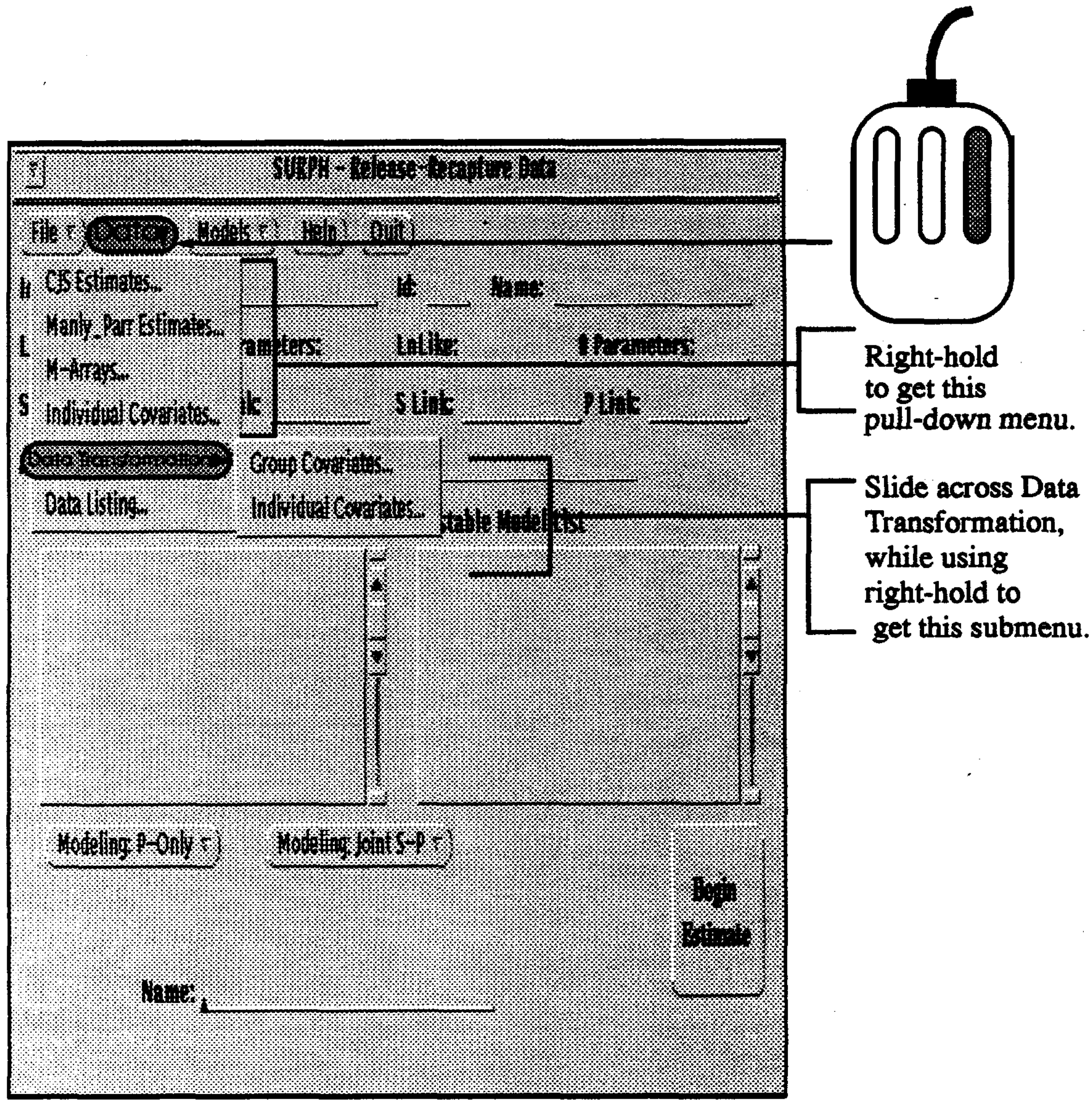
Product of Final Period Survival/Capture

Population	
1	0.427 (0.032)
2	0.467 (0.039)
3	0.301 (0.040)
4	0.282 (0.038)
5	0.177 (0.039)
6	0.156 (0.038)

Text windows generally consist of the Text Pane, where the information is displayed, and a small "control panel", which **in** most **cases** contains a **Print** button and a **Quit** button.

When you quit SURPH, all information associated with a particular data set is no longer accessible. At times, you may wish to retain the information within the text window for further inspection after you have exited SURPH. To learn how to do so, see

Storing Text Pane Information .



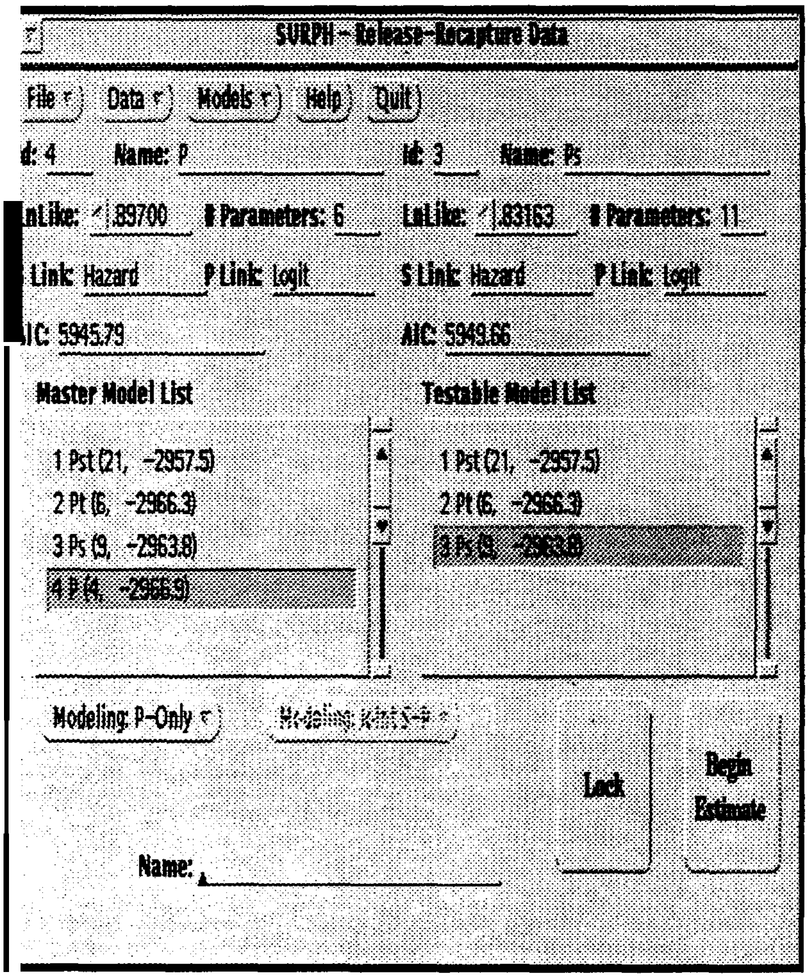
This is the pull-down and pull-right menu the user gets if he accesses the **Data Button and the Data Transformation** menu option.

Right-hold to get the pull-down.

Move the mouse down to select **Data Transformation,** then move sideways to the right, while holding the right mouse key down to get the pull-right menu. Release the right mouse key when the cursor covers the desired option (i.e., Group Covariates or Individual Covariates).

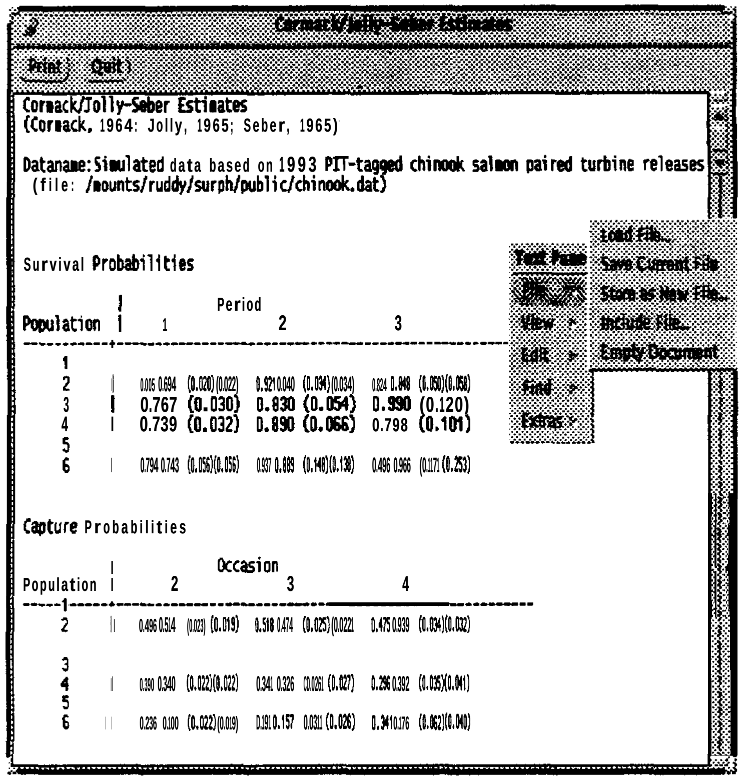
Alternatively, you may right-click on the **Data Button,** right-click on the **Data Transformation** option in pull-down, then left-click on the desired option in the pull-right menu.

THIS PAGE INTENTIONALLY LEFT BLANK



This graphic shows the typical appearance of the SURPH Base Window after you estimate a series of models for capture probabilities **under Modeling: P-Only. After** a series of capture models is estimated, you have the opportunity to choose which capture model will be retained for the survival modeling phase of **the** data analysis. All models that have been estimated during the current SURPH session, and those that have been loaded f'rom previously-saved SURPH model files, appear within the **Master Model List.** For a model to appear in the **Testable Model List,** one of two conditions must be met. Either the selected model within the **Master Model List is** nested within the model in the **Testable Model List** or the model within the **Testable Model Lii is nested within the** selected model within the **Master Model List. Two** models are said to be "nested", or "hierarchical", when one model is a special case of **the** other. For example, Model A is nested in Model B if Model A can be derived from Model B by either equating two or more parameters of B or by simply dropping'one or more parameters of B.

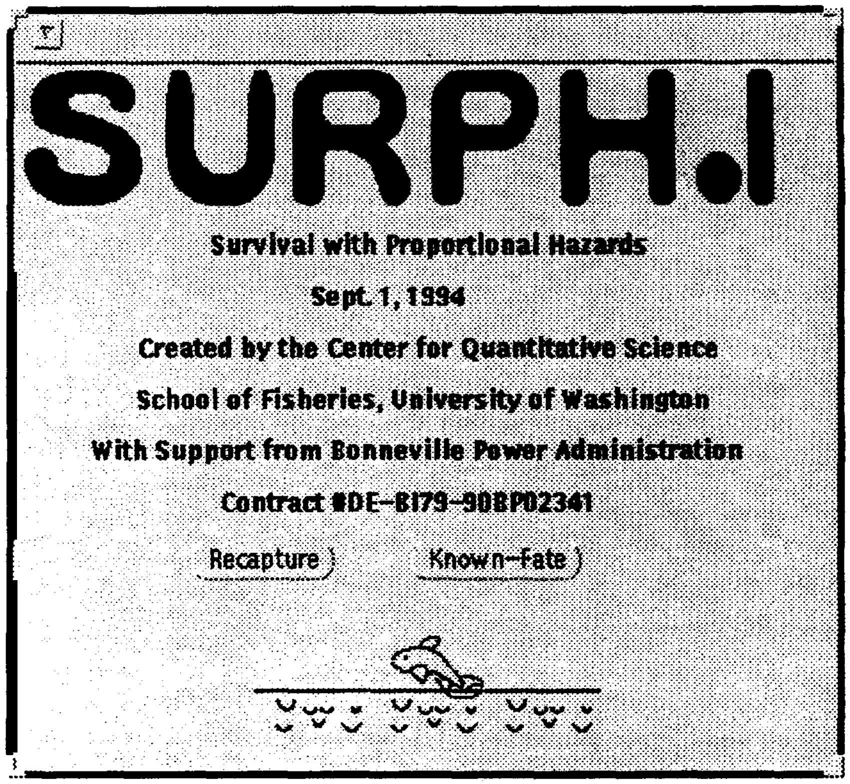
To decide which model best fits the data, SURPH allows you to compare nested models using the Likelihood Ratio Test **(LRT) (Wald 1943),** and to compare non-nested models using Akaike's Information Criterion (AK!) (Akaike 1973). To test nested models using the LRT, select a model from **the** Master Model List. Nested models will appear within the **Testable Model List.** When you select a model in the **Testable Model List,** a pop-up window will appear. This pop-up window displays the change in Deviance between the two models (Test Statistic), the difference in **the** number of parameters estimated between the two models **(df),** and the probability of detecting a change in Deviance of such magnitude if the more complex model does not improve the **fit** (p-value).



To write text from a text pane to a file (e.g., the output from **Data -> Manly-Parr),** use the mouse to move the pointer onto the text pane, and right-click. This will cause the **Text Pane Menu** to appear. Right-click on **"File"** to bring up the pull-right menu. Then left-click on the **"Store** as New **File"** option. A **Text: Store** Pop-Up Window will appear. You can now store the file in a text file that can be **editted** using a text editor, and printed using standard UNIX utilities.

Occasionally, the display within the text window is too wide to be seen without enlarging the window. If this is the case, then the display is typically too wide to be printed using **the Print Button.** To print the entire contents of the display, you must save the output text to a new text file, import the new text file into a text editor, and print the file from the text editor. Even when the information fits comfortably within the text pane, and **can be** printed **in** its **entirity** using the **Print Button, this** technique can be used to save the text for future reference.

THIS PAGE INTENTIONALLY LEFT BLANK



To begin a SURPH session, type either

> surph *datafile*

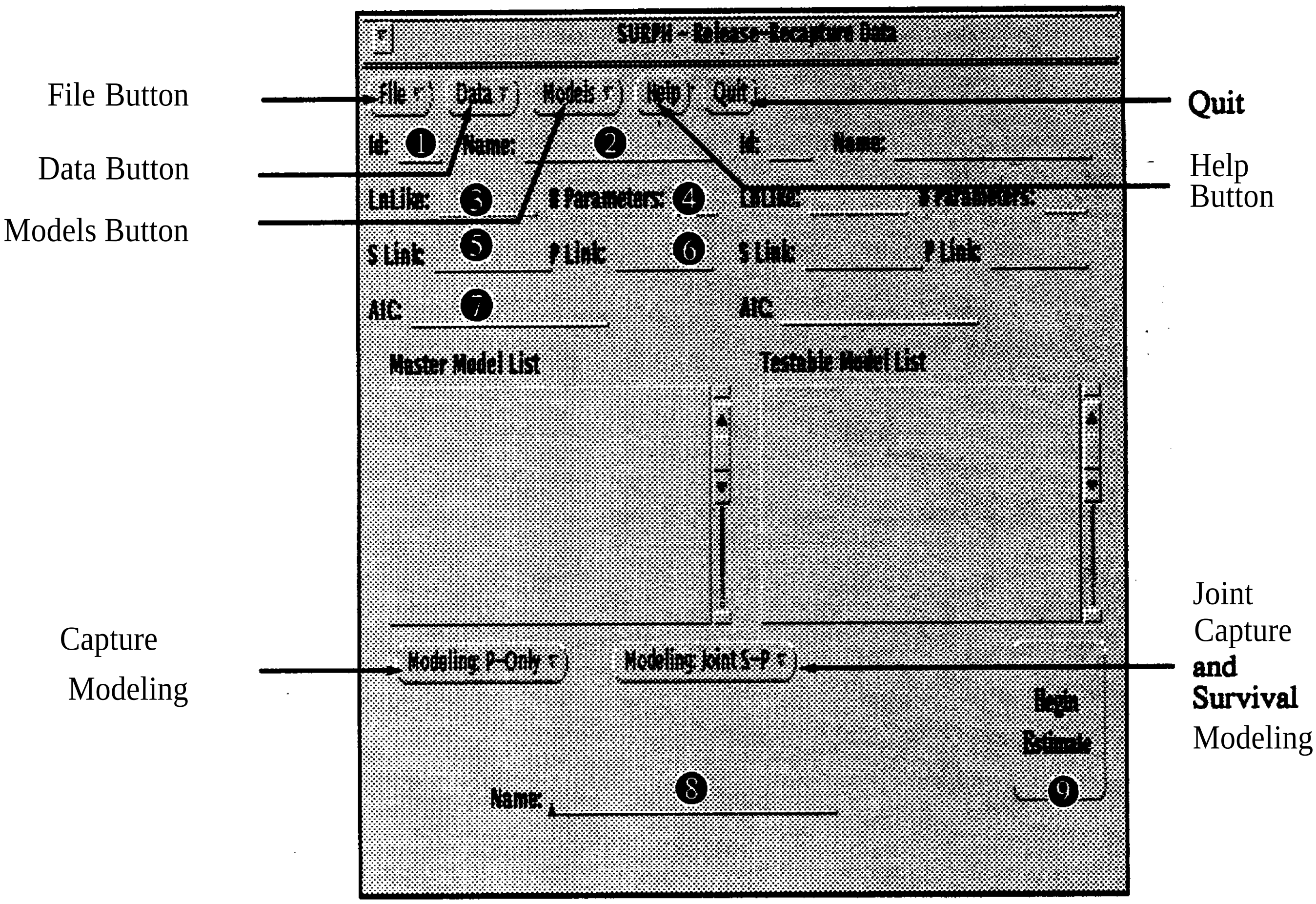
or

> surph

In **the** first case, the user **specifies** the **datafile** that will be used in the analysis on the command line. In the second case, the user does not specify the dataiile at this time; the **datafile** to be used in the analysis will be specified once the **SURPH** Base **Window** appears **(See SURPH Base Window** and **File: Load Data entries in** Chapter 7 of the **SURPH**. 1 manual).

To continue, the user must select whether the data are from a Release-Recapture study **(Recapture Button)** or a Known-Fate study **(Known Fate Button). Once** the appropriate button is selected, the **SURPH** Base Window will appear.

THIS PAGE INTENTIONALLY LEFT BLANK



This is the SURPH Base Window for Release-Recapture data files, the first window that appears after you select the Release-Recapture modeling option from the SURPH Introductory Window. The SURPH Introductory Window appears when the program is started with a command line:

> surph ***datafile***

All data summaries, modeling operations, **graphics,** and diagnostics are accessed from the SURPH Base Window.

Overview of Command Buttons

The pull-down menu accessed by the **File Button** is used to load data files, and to load or store model files. The **File** pull-down can also **be** used to end the SURPH session. Because the data file was specified on the command line, the Load Data command **(File -> Load Data)** is disabled.

The pull-down menu **accessed** by **the Data Button has** choices for graphical data display; computed Cormack **(1964),** Jolly **(1965),** Seber (1965) survival and capture probability estimates and Manly-Parr (Manly and **Parr** 1968) capture probability estimates; summary statistics in the form of M-Arrays **(Burnham** et al. 1987); data transformations; and echoing input data to the screen.

The pull-down menu **accessed by the Models Button** allows display of model estimates and graphical displays of model fit. This button also allows you to discard models and change model names.

The Help Button explains the use of the various buttons. To see further instructions, left-click on this button.

The Quit Button ends the SURPH session.

The Modeling: P-Only Button is used to **parameterize** the capture-only model. When this button is used, the capture process is **modelled** as an extension of the Manly-Parr (1968) Model and uses only a subset of the total release data. Menu choices under the **Modeling: P-Only Button allow you to** model the capture process at the population or individual level, and allow you to specify the link function. For further descriptions of the functions associated with this button, see Capture Modeling .

The Modeling: Joint S-P Button is used to **parameterize** the joint survival/capture model. When this button is used, the capture process and survival process are **modelled** jointly. Menu choices under **the Modeling: Joint S-P Button allow** you to model the capture and survival processes at the population or individual level, and allow you to specify the link functions for both. The final period product of capture and survival can also be **modelled** at the population or individual level, and you can specify the link function. For further descriptions of the functions associated with this button, see Joint Modeling: Pull-Down Menu.

The SURPH Base Window **has two lists** of estimated models, **the Master Model List** and the **Testable Model List. The Master Model List** contains all models that have been estimated during the current SURPH session, and those that have been loaded from previous SURPH sessions. The **Testable Model List** contains models that can be compared to the model that has been selected from the **Master Model List.** For a **model to** appear **within the** Testable Model **List,** either the model within the **Testable** Model List is nested within the model selected **from the Master Model List,** or the model selected from the **Master Model List** is nested **within the** model **in the Testable Model List** (See Testing Nested Models: Master and Testable Lists).

Numbered Features

The numbered spaces 1-7 refer to attributes of the model selected from the Master **Model List.** Analogous spaces in the upper right-hand portion of the SURPH Base Window refer to attributes of **the** model selected **within the Testable Model List.**

1 Displays the Model ID Number of the selected model **in the Master Model List.**

2 Displays the name of the selected model in the **Master Model List.**

3 Displays the Log-Likelihood of the selected model **in the Master Model List.**

4 Displays the number of parameters estimated **in the** selected model **in the Master Model List.**

5 Displays the link function for the survival parameters of the selected model in the **Master Model List.**

6 Displays the link function for the capture parameters of the selected model in the **Muster Model List.**

7 Displays Akaike's Information Criterion (Akaike 1973) **-** This is used to help determine the best fitting models when the models are not nested.

The space numbered 8, and the button numbered 9 are used to create new models:

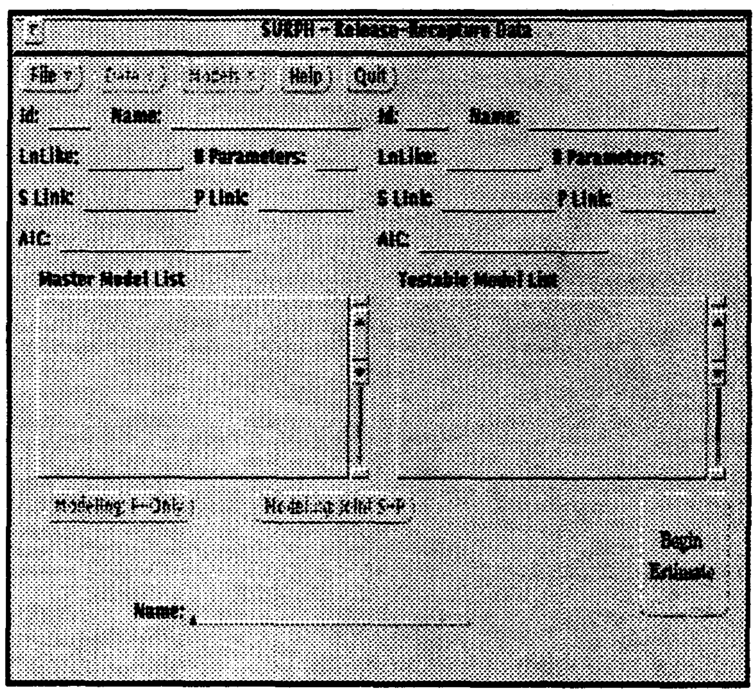
8 This field allows you to name the model prior to running the estimation procedure. If you does not specify a name for the model, the name "Model 'n"' is assigned, where "n" is the Model ID Number. If you have a model you wish to rename, **see the** section for the **Models Button.**

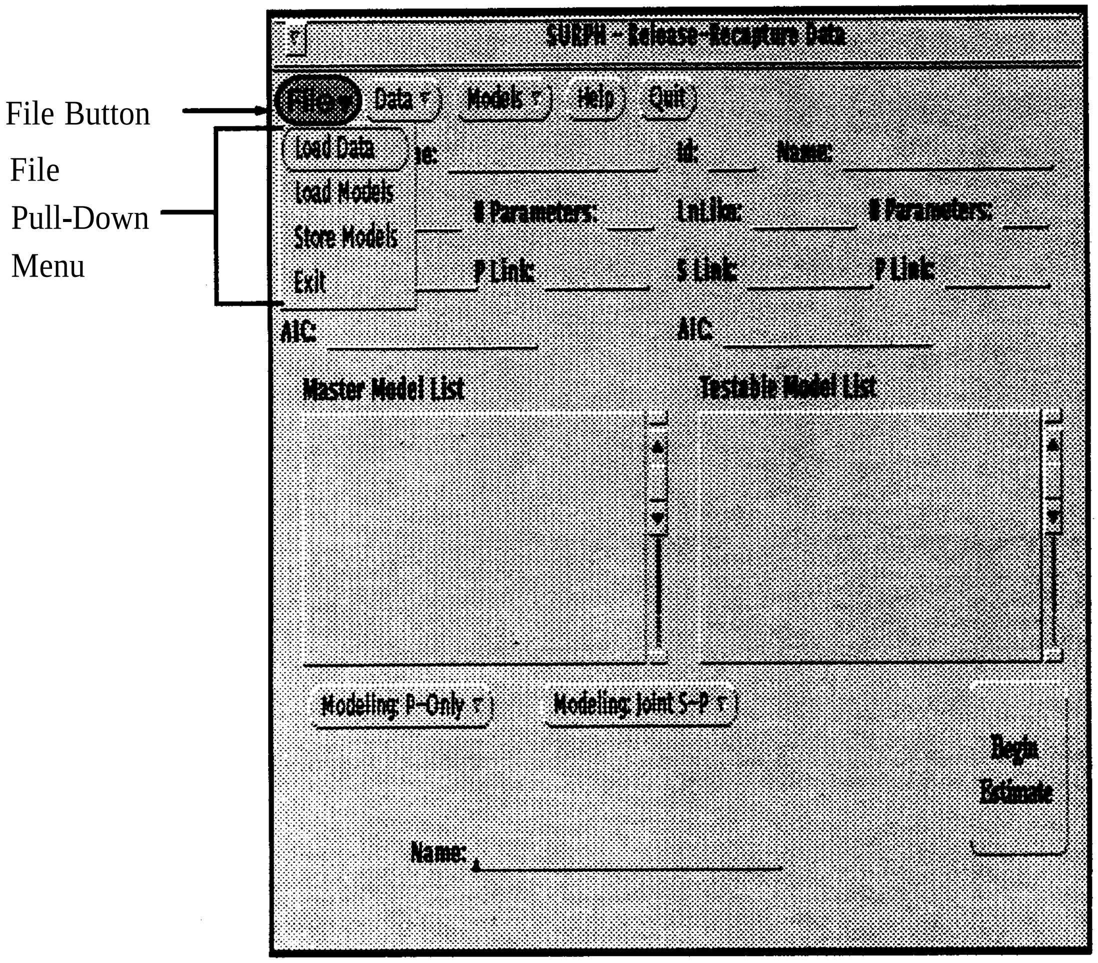
9 Left-clicking on this button initiates the numerical optimization routine to estimate **the** parameters of the model specified **in the** modeling options under **Modeling: P-Only** or **Modeling: Joint S-P.**

If you start **SURPH** without specifying an input data file on the command line:

> surph

not all buttons are activated (see below). Before these de-activated **buttons** can be used, data must **be** loaded. To load **data,** right-click on **the File Button.** A pull-down menu **will** appear. Left-click on **Load Data. The Load Input File** window will appear. If your data file is in the directory **from** which you started SURPH, left-click **in the File** text field and type the name of the **file** that contains your data. If your data is in a different directory, left-click in the **Directory** text field and change the **directory** path in the **Directory field,** then left-click in the **File** field and enter the file name in the blank **File field.** L&t-click on the **Load Button in the Load Input Data** window to load the data.





To access this pull-down menu, right-click on the **File Button.**

The commands available from this pull-down menu are:

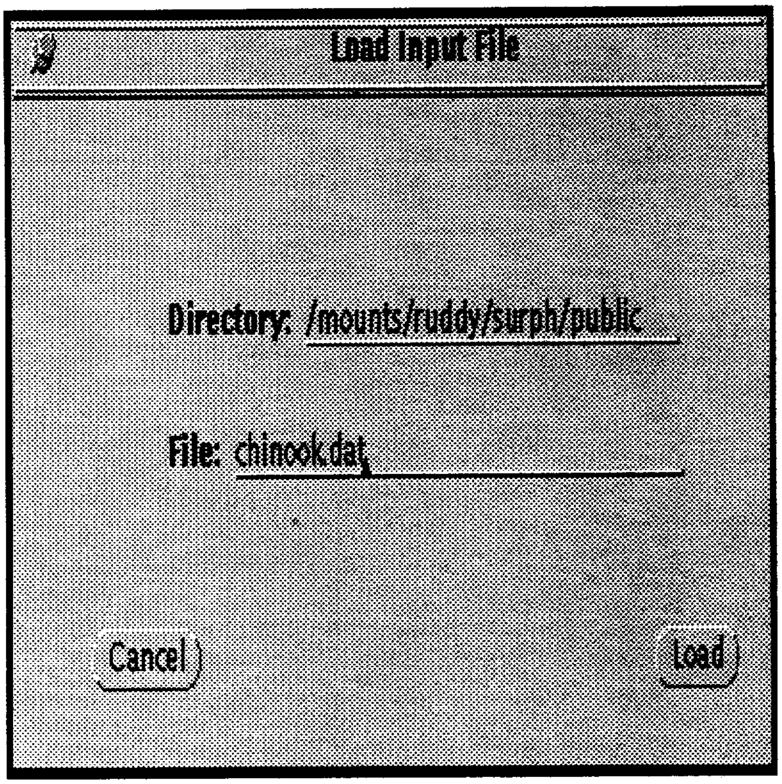
Load Data - **Load** a data file to be analyzed (See **Load Data**).

Load Models - **Load** previously developed and stored models (See **Load Models**).

Store Models - Save models from the **Master Model List** for future retrieval (See **Store Models**).

Exit - Quit SURPH

THIS PAGE INTENTIONALLY LEFT BLANK



This pop-up window **(File -> Load Data)** is used to load data for SURPH. Alternatively, the data file can be specified on the UNIX command line by typing:

> surph *datafile*

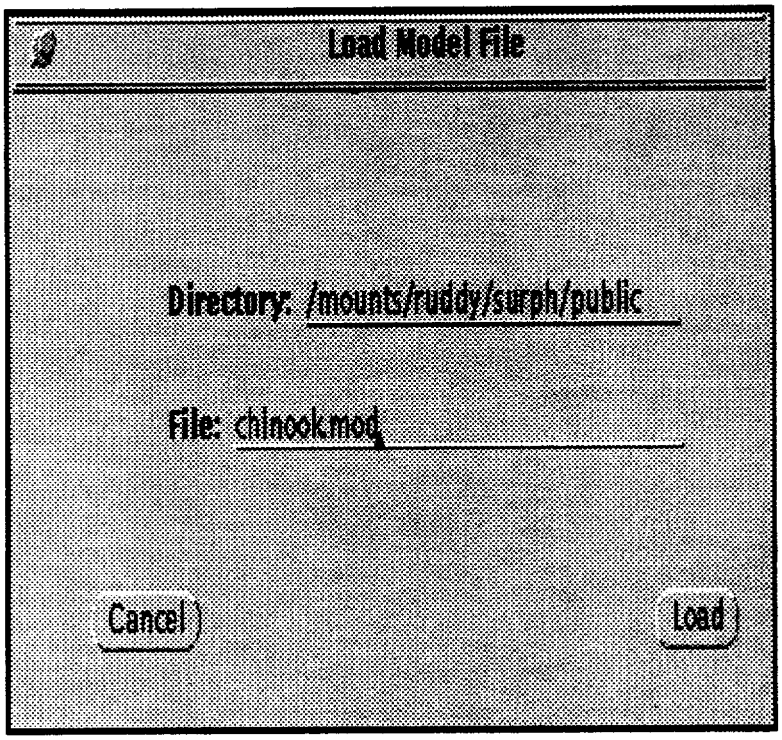
Data files for SURPH consist of definitions of the dimensions of the data (e.g., number of populations), followed by recapture data for each tagged animal. Examples of data files can be found in Chapter 6 of the SURPH. 1 manual.

To specify the directory where the **data** file **is** located, left-click on **the Directory** field, and type the directory path. To specify the data file, left-click on **the File** field, and type the data file name. Once the correct directory has been specified and the data file name has been entered in the **File** field, initiate loading by left-clicking on the **Load Button.**

In this example, the file **"chinook.dat"** in **the** directory **"/mounts/ruddy/surph/** public" would be loaded.

Left-click on the **Cancel Button** to dismiss this window without loading data.

THIS PAGE INTENTIONALLY LEFT BLANK



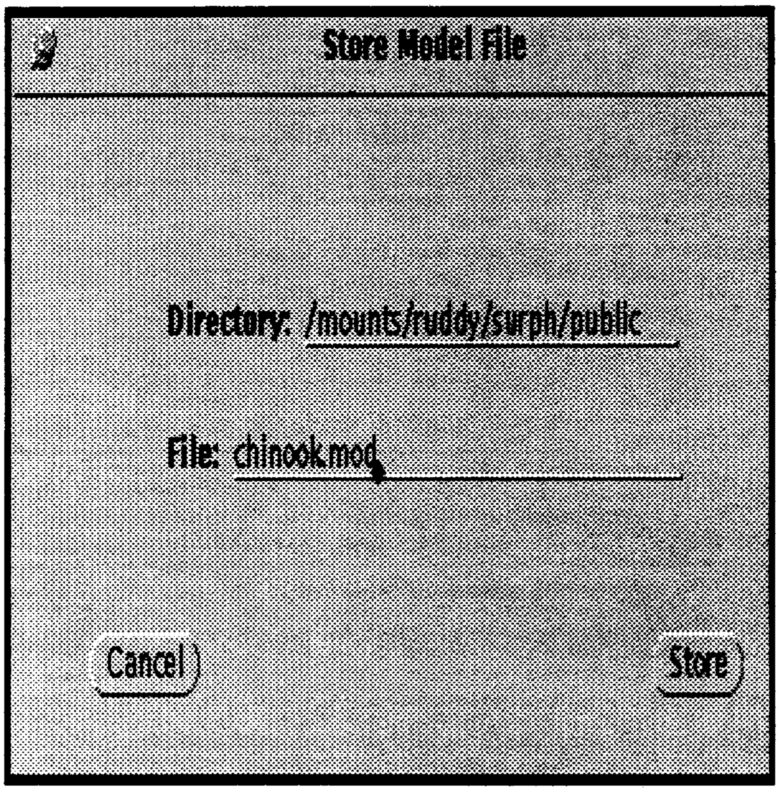
Tbis pop-up window **(File -> Load Model)** allows you to load previously stored SURPH models **into the Master Model List. Models** must have been saved using SURPH during a previous modeling session. Model files are formatted ASCII files that contain all pertinent information regarding the stored SURPH models.

To specify the directory where the model file **is** located, left-click on **the Directory** field, and type the correct directory path. To specify the file that contains the previously stored models, left-click on the **File** field, and type the name of the model file. Once the correct directory has been specified and the name of the model file has been entered **into the File** field, left-click on the **Load Button** to load models.

In this example, the file **"chinook.mod"** in the directory **"/mounts/ruddy/surph/**public" would be loaded. **"Chinook.mod"** holds models that were saved during a previous **SURPH** session. After loading, these models would be accessible to you in SURPH, just as if they had been created in the current session.

Left-click on the **Cancel Button to dismiss this** window without loading models.

THIS PAGE INTENTIONALLY LEFT BLANK



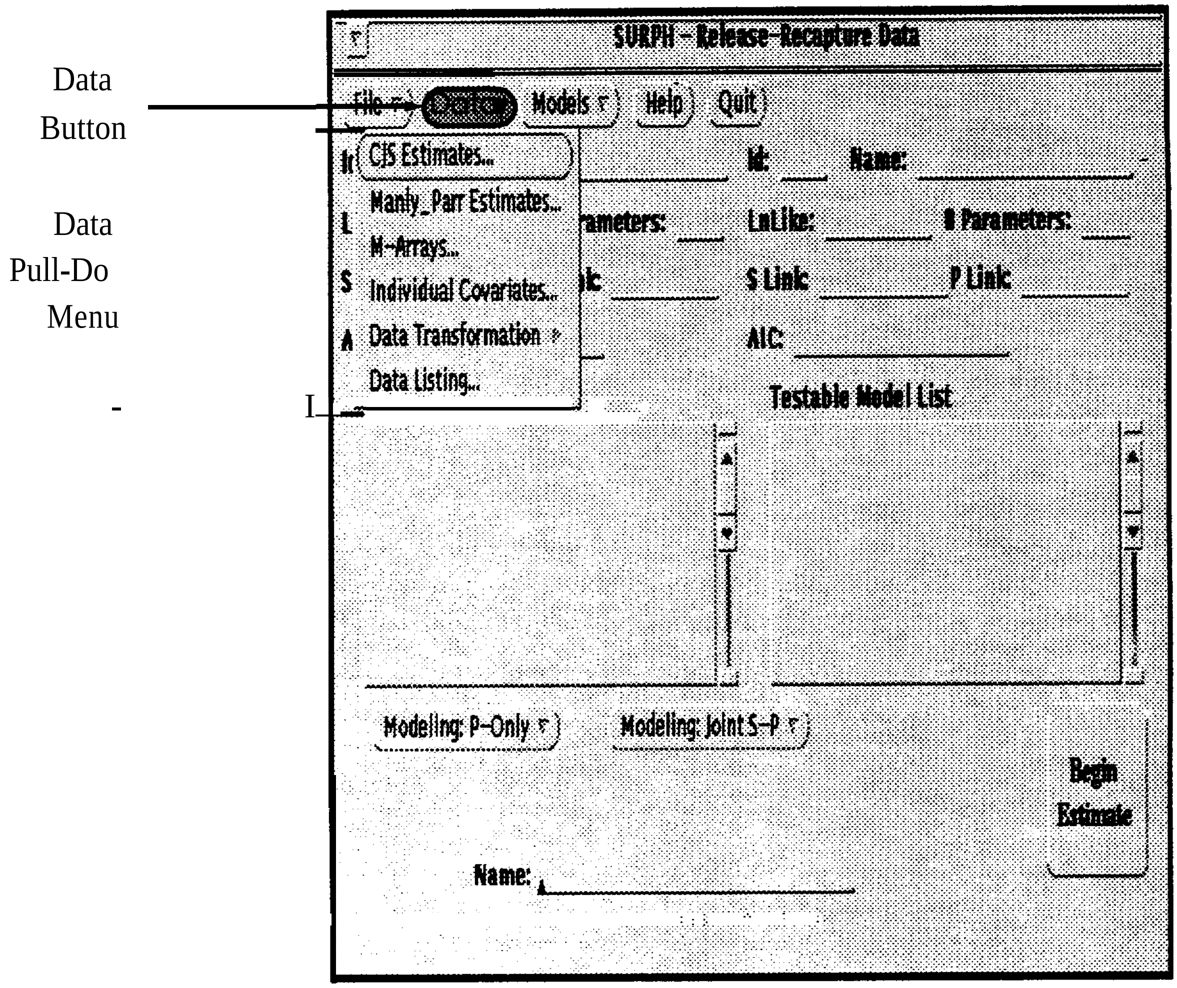
This pop-up window **(File -> Store Model) allows you to** store SURPH models from **the Master Model List** to a disk file. Model files are formatted ASCII, containing all pertinent information regarding the **SURPH** models.

To specify the directory where the models are located, left-click on **the Directory** field, and type the directory path. To specify the tile where the models are located, left-click on the **File** field, and type the name of the model file. Once the correct directory has been specified and the file name has been entered into the field, left-click on the **Store Button** to store the models.

In tbis example, the model file **"chinook.mod"** would be created in the directory **"/mounts/ruddy/surph/public"**. **"Chinook.mod"** would hold all models that had been estimated during the current SURPH session. After the models have been stored, these models can be loaded during future SURPH sessions.

Left-click on tbe **Cancel Button** to dismiss this window without storing models.

THIS PAGE INTENTIONALLY LEFT BLANK



To access this pull-down menu, right-click on the **Data Button.**

The commands available from this menu produce data summaries and transformations:

CJS Estimates - Provides Cormack (1964), Jolly (1965), Seber (1965) estimates of survival and capture probabilities. For further information, see Data: Cormack/Jolly-Seber Estimates .

Manly-Parr - Provides estimates of capture probabilities using the Manly-Parr (1968) method, which is independent of the survival process. For further information, see Data : Manly Parr .

M-Arrays Provides release-recapture summary statistics (Burnham et al. 1987). For further information,

see Data : M-Arrays .

Individual Covariates - Displays graphical representations (histograms and cumulative distribution functions) of the data, by covariate, by population, by period. For further information,

see Individual Covariates - Histogram or Individual Covariates - CDF Plot .

Data Transformation - Allows transformations of covariate data. Examples of transformations are squares, square roots, cross products, reciprocals and natural logarithms. For further information,

see Data Transformation : Group Covariates or Data Transformation : Individual Covariates

Data Listing Gives a complete listing of the input data in a formatted display. For further information,

see Data Listing .

Cormack/Jolly-Seber Estimates

Print Quit

Cormack/Jolly-Seber Estimates
(Cormack, 1964; Jolly, 1965; Seber, 1965)

Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
(file: /mounts/ruddy/surph/public/chinook.dat)

Survival Probabilities

Population	Period 1	2	3
1	0.885 (0.020)	0.921 (0.034)	0.824 (0.050)
2	0.684 (0.022)	0.840 (0.034)	0.848 (0.058)
3	0.767 (0.030)	0.830 (0.054)	0.990 (0.1201
4	0.739 (0.032)	0.890 (0.066)	0.798 (0.101)
5	0.794 (0.056)	0.937 (0.148)	0.4% (0.117)
6	0.743 (0.056)	0.889 (0.138)	0.966 (0.253)

Capture Probabilities

Population	Occasion 2	3	4
1	0.514 (0.019)	0.474 (0.022)	0.539 (0.032)
2	0.4% (0.023)	0.518 (0.025)	0.475 (0.034)
3	0.390 (0.022)	0.341 (0.026)	0.296 (0.035)
4	0.348 (0.022)	0.326 (0.027)	0.352 (0.041)
5	0.2% (0.022)	0.191 (0.031)	0.341 (0.062)
6	0.180 (0.019)	0.157 (0.026)	0.176 (0.040)

Product of Final Period Survival/Capture

Population	
1	0.427 (0.032)
2	0.467 (0.039)
3	0.301 (0.040)
4	0.262 (0.038)
5	0.177 (0.039)
6	0.156 (0.038)

This text pane **(Data -> CJS Estimates) shows the** Cormack **(1964),** Jolly **(1965),** Seber (1965) survival and capture probability estimates. The parenthetical number beside each probability estimate is the estimated standard error for that parameter. For example, for Population 1 in Period 1, the estimated survival probability is 0.885, with an estimated standard error of 0.020, while the estimate of the product of final period survival and capture probabilities for Population 3 is 0.301, with an estimated standard error of 0.040.

THIS PAGE INTENTIONALLY LEFT BLANK

```
Manly-Parr Estimates
Print   Quit

Manly-Parr Capture Probability Estimates
(Manly and Parr, 1968)

Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
 (File: /mounts/ruddy/surph/public/chinook.dat)


Unadjusted for Removals
           |                   Occasion
Population |        2               3               4
-----------+------------------------------------------------
         1 |   0.422(229/543)  0.387(185/478)  0.439(100/228)
         2 |   0.405(160/395)  0.444(163/367)  0.384( 78/203)
         3 |   0.298(111/372)  0.269( 94/350)  0.215( 40/186)
         4 |   0.265( 95/358)  0.259( 81/313)  0.272( 40/147)
         5 |   0.184( 48/261)  0.138( 27/196)  0.254( 17/67 )
         6 |   0.120( 30/249)  0.114( 27/237)  0.123( 14/114)

Removal Proportions
           |                   Occasio
Population |        2               3               4
-----------+------------------------------------------------
         1 |   0.303(138/455)  0.299(130/435)  0.331(116/350)
         2 |   0.319(108/339)  0.256( 92/359)  0.310( 75/242)
         3 |   0.284( 85/299)  0.292( 77/264)  0.348( 71/204)
         4 |   0.319( 82/257)  0.278( 73/263)  0.311 ( 64/206)
         5 |   0.326( 61/187)  0.324( 57/176)  0.342( 50/146)
         6 |   0.291( 39/134)  0.308( 41/133)  0.343( 47/137)

Manly-Parr estimates adjusted for removals
           |                   Occasion
Population |        2               3               4
-----------+------------------------------------------------
         1 |
         2 |   0.511 0.500 0.373  0.474 0.518 0.341  0.539 0.475
         3 |                                     0.296
         4 |     0.347            0.326           0.352
         5 |                                     .....
         6 |   0.251 0.162       0.191 0.157     0.341 0.176
```

This text pane displays Manly-Parr (1968) capture probability estimates **(Data -> Manly-Parr)** for each sampling occasion and population. The **first** table displays the uncorrected point estimates of the capture probabilities and the number of captures used in their calculations. The formulas for the capture probabilities are given in Section 4.6. For example, in Population 1,478 animals were caught both before and after the **3rd** sampling occasion, and of those, 185 were also captured on the **3rd** sampling occasion, for a capture probability estimate of **185/478** = 0.387. These estimates are biased when trap mortality or known removal of captured animals occurs.

The second table displays the proportion and fraction of the captured animals that were removed from each population on each sampling occasion. For example, in Population 1,435 animals were caught on the **3rd** sampling occasion, and of those, 130 were either killed or removed, for a removal proportion of **130/435** = 0.299.

The final table displays the estimated capture probabilties corrected for trap mortality and known removals. For an explanation of the **adjustment** procedure, see Section 4.6 in the **SURPH.1** manual.

THIS PAGE INTENTIONALLY LEFT BLANK

Reduced M-Arrays

Print) Quit)

Data Summary: Reduced M-Arrays
(Burnham, et al, 1987; pp. 32-37)

Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
(file: /mounts/ruddy/surph/public/chinook.dat)

Population 1

❶ Occasion	Released ❷ Total	❸ Occasion Next Captured 2	3	4	5	❹ Never Recaptured
1	1000	455:138	190:54	80:25	44:0	231
2	567		245:76	128:48	41:0	153
3	305			142:43	43:0	120
4	234				100:0	134

Population 2

Occasion	Released Total	Occasion Next Captured 2	3	4	5	Never Recaptured
1						
2	1000	339:108	134:38	64:20	35:039:0	426 153
3	481'		225:54	112:34	51:0	104
4	167				78:0	89

Population 3

Occasion	Released Total	Occasion Next Captured 2	3	4	5	Never Recaptured
1	1000	299:85	137:41	73:25	51:0	440
2	464		127:36	77:33	55:0	205
3	187			54:13	40:0	93
4	133				40:0	93

This text pane **(Data -> M-Arrays)** displays numbers of animals released and subsequently recaptured, in the format of the "reduced M-array" of **Burnham** et al. (1987). There is one table for each population. The elements of the table for each population are:

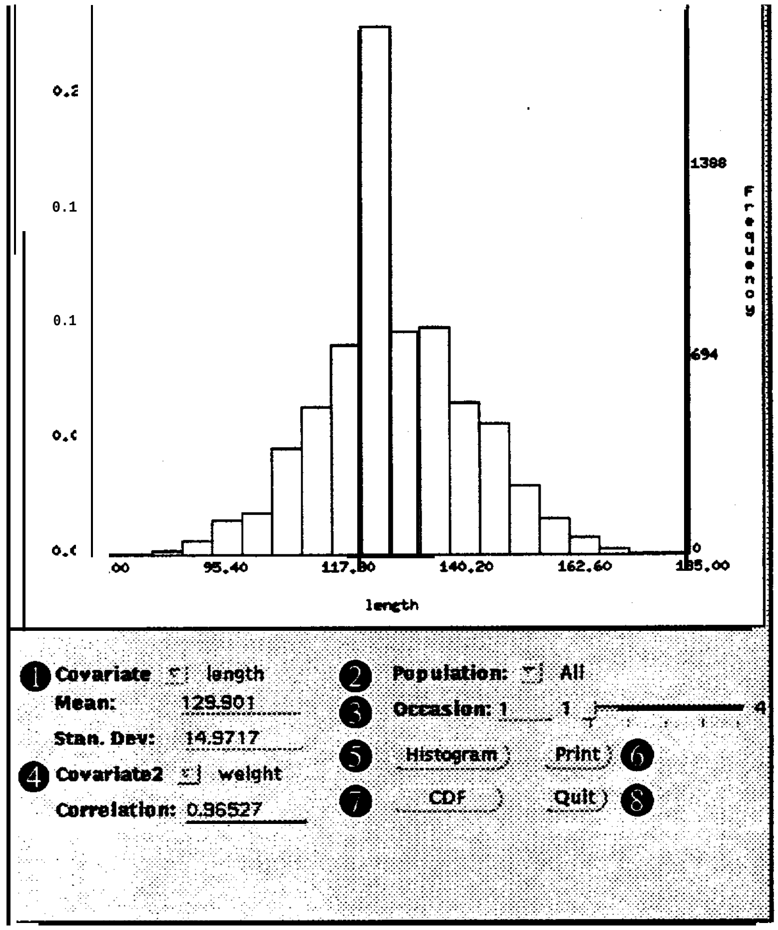
❶ Release Occasion — Occasion of release of marked animals.

❷ Total Released — Number of animals seen and released at this occasion. This total includes animals marked and released for the **first** time, as well as previously-marked animals recaptured on this sampling occasion, and re-released.

❸ Occasion Next Captured — For each release occasion, the number of animals first recaptured again on each subsequent occasion. The number to the left of the. colon is the total number of animals captured. To the right of the colon is the number of animals not re-released, due either to trap mortality or other known removal.

❹ Never Recaptured — For each release occasion, the number of animals that were never again seen after the initial release.

In the example shown, 1000 animals in Population 1 were marked and released on occasion 1. Of these, 455 were recaptured. on the 2nd sampling occasion. Of the 545 animals not seen on occasion 2,190 were &captured on occasion 3. Of the 455 **recaptured** on occasion 2,138 were removed. The other 317 were combined with 250 newly marked animals on occasion 2 for a total of 567 released on occasion 2. A total of **435** animals from Population 1 were recaptured on occasion 3 (190 + **245),** and of those 130 (54 + 76) were removed. The remaining 305 were released on occasion 3, with no newly marked **animals**.



This graphic **(Data -> Individual Covariates)** displays histograms of individual covariate data. Histograms display individual covariate data from all animals from the selected population known to be alive on the selected occasion. Different individual covariates may be selected using the covariate pull-down menu **(❶)**. Similarly, the different populations may be selected using the population **pull-down** menu **(❷)**. To access either, right-click on the small arrow adjacent to the appropriate keyword **(Covariate** or **Population).**

Numbered Features

0 Covariate Pull-Down Menu. This menu allows you to select the individual covariate to be viewed. Right-click on the downward pointing arrow to access the menu. Left-click on a menu choice to select a covariate.

❷ Population Pull-Down Menu. This menu allows you to select the population to be viewed Right-click on the downward pointing arrow to access the menu. **Left-click on** a menu choice to select a population. The menu includes a choice to view data for all populations simultaneously.

❸ Occasion Designation. This slide-bar allows you to select the occasion to be viewed. Pull slide bar (left-hold on the small rectangle on the slide bar and move mouse sideways) across to increment or decrement the occasion designation.

❹ **Covariate2** Pull-Down Menu - This menu allows you to select a second individual covariate to use in conjunction with the individual covariate selected above (❶) in computing the simple correlation (see below). Right-click on the downward pointing arrow to access the menu. Left-click on a menu choice to select a covariate.

❺ **Histogram Button.** Displays the data as a histogram (as shown here).

❻ **Print Button.** Sends a copy of the screen to the printer.

❼ **CDF Button.** Displays the cumulative distribution of covariate values for recaptured and non-recaptured animals. The cumulative distribution plot is explained under Individual Covariates - CDF Plot .

❽ **Quit Button.** Quits this window.

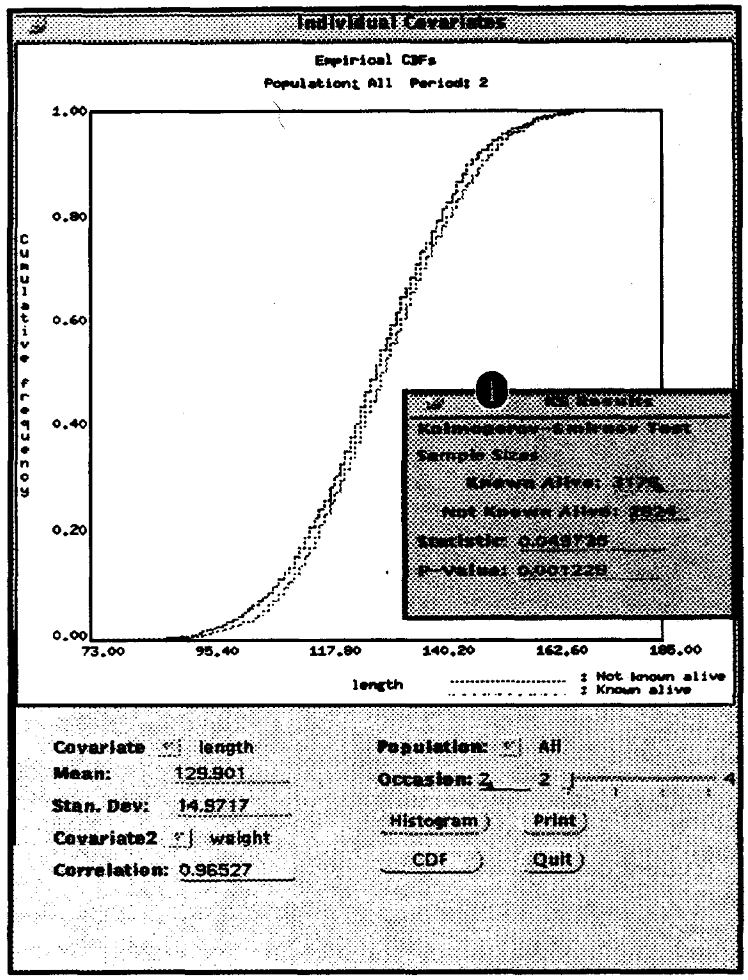
Statistical Functions:

Mean - Displays the arithmetic mean for the selected covariate for the selected population. Data from all animals in the population, regardless of marking occasion, are combined for this calculation. In this example data set, we have designated **"length"** as our primary covariate. The mean length for all animals released was 129.901.

Stand. Dev. **-** Displays the standard deviation for the covariate and the population currently selected. Data from all animals in the population, regardless of marking occasion, are combined for this calculation. In this example data set, we have designated "length" as our primary covariate. The standard deviation of the length for all animals released was 14.9717.

Correlation - Displays the simple correlation (Zar 1984) between the covariate selected from the **Covariate pulldown** menu and the covariate selected from the **Covariate2 pulldown** menu. Data from all animals in the population, regardless of marking occasion, are combined for this calculation. In the example, we have designated "length" as our primary covariate and "weight" as Covariate2. The correlation between "length" and "weight" in this example data set is 0.96.

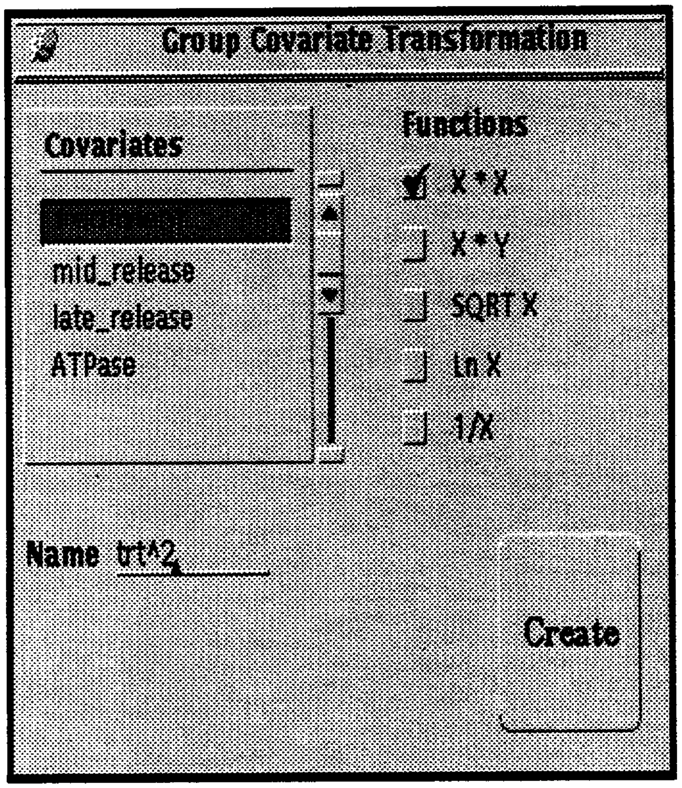
THIS PAGE INTENTIONALLY LEFT BLANK



This graphic **(Data -> Individual Covariates)** displays cumulative distribution functions (cdfs) of individual covariate data. **Two** cumulative distributions are displayed: one for previously marked animals that were known to be alive on the selected occasion (recaptured on the selected occasion or later), and one for those not known to be alive (captured for the last time prior to the selected occasion). Large differences between the curves may indicate some selective effect of the individual factor on capture, survival, or both.

For further information concerning other functions in this window, see Individual Covariates: Histogram .

In addition to the graphic display of cdfs, the Kolmogorov-Smirnov test for homogeneity of discrete distributions is automatically computed. The K-S pop-up window (❶) displays the number of animals known alive, the number of animals not known alive, the maximum difference between **the** two empirical distributions (Statistic), and the probability (P-Value) of observing such a difference if the two samples were from identical distributions.



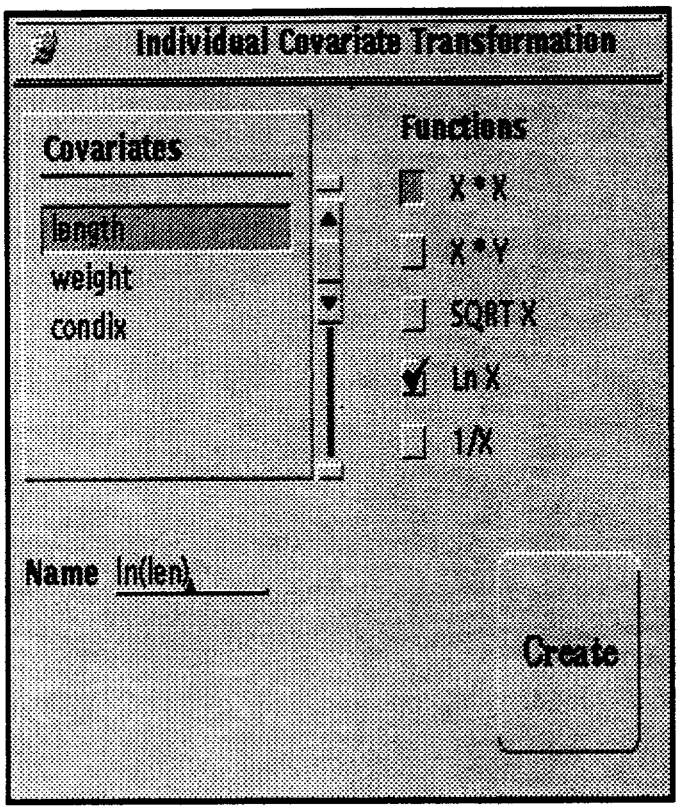
This window **(Data -> Data Transformations -> Group Covariates) is used to** generate new Group Covariates that are functions **of** the currently **defined** Group Covariates. To create new data that are a transformation of current data, select the covariate(s) from the Covariates List and the functional transformation from the Function List. To name the new Covariate, type the name you desire in the blank line following "Name" prior to pushing **the Create Button.** If you forget to include a name for the new variable, a pop-up window will appear and advise you to name the new variable. You cannot create a new variable without naming it. To create the new Covariate, place the **arrow** key above the **Create Button,** then left-click.

Available functional manipulations are:

X*X	Square the value of the selected Covariate.
SQRT (X)	Take the Square Root of the value of the selected Covariate
Ln (X)	Take the Natural Logarithm of the value of the selected Covariate.
1/ (X)	Take the Reciprocal of the value of the selected Covariate

X*Y Multiply the value of one selected Covariate by the value of another **selected** Covariate

In the example window, left-clicking on **the Create Button will create** a new variable ***"trt^2"***. ***The values*** of ***trt^2*** for each population will be equal to the square of the ***values*** for ***treatment.***



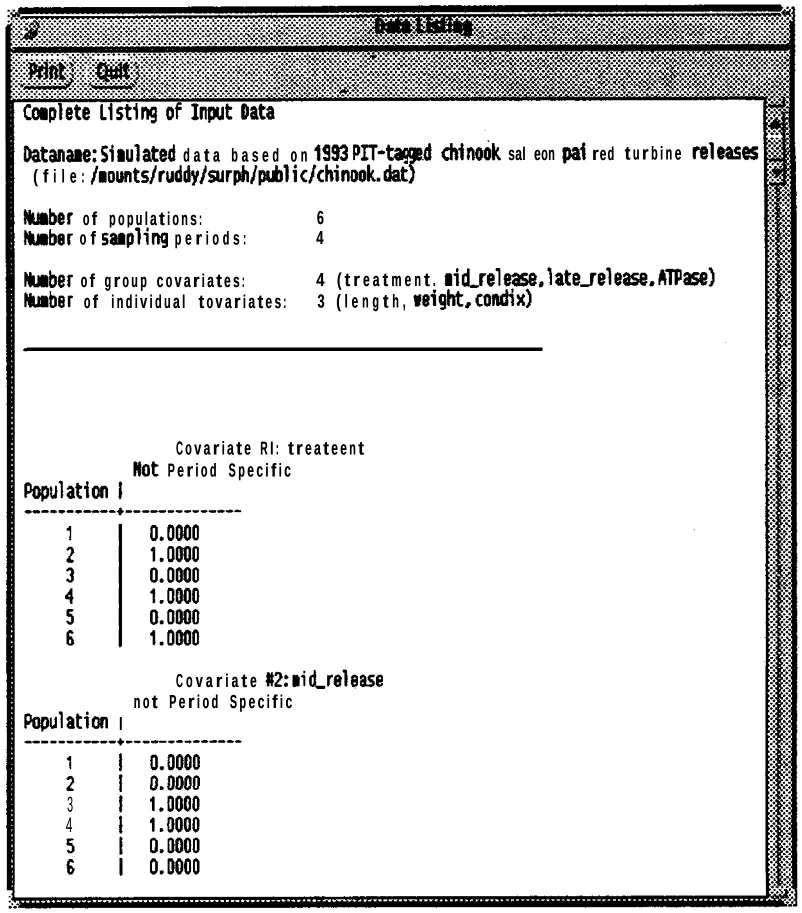
This window **(Data -> Data Transformations -> Individual Covariates) is used** to generate new Individual Covariates that are functions of the currently selected Individual Covariate data. To create new data that are a transformation of current data, select the **covariate(s) from** the Covariates List and the functional transformation from the Function List. To name the new Covariate, type the name you desire in the blank line following "Name" prior to pushing the **Create Button.** If you forget to include a name for the new variable, a pop-up window will appear and advise you to name the new variable. You cannot **creat** a new variable without naming it. To create the new Covariate, place the **arrow** key above the **Create Button,** then left-click.

Available functional manipulations are:

x*x	Square the value of the selected Covariate.
SQRT (X)	Take the Square Root of the value of the **selected** Covariate.
Ln (X)	Take the Natural Logarithm of the value of the selected Covariate.
1/ (X)	Take the Reciprocal of the value of the selected Covariate.

X*Y Multiply the value of one selected Covariate by the value of another selected **Covariate.**

In the example window, left-click on **the Create Button to create** a new variable ***"ln(len)"***. The values in ***ln(len)*** will be equal to the Natural **Logarithm** of the values in ***length.***



Data Listing

Print Quit

Complete Listing of Input Data

Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
(file: /mounts/ruddy/surph/public/chinook.dat)

Number of populations: 6
Number of sampling periods: 4

Number of group covariates: 4 (treatment, mid_release, late_release, ATPase)
Number of individual covariates: 3 (length, weight, condix)

Covariate #1: treatment
Not Period Specific

Population	
1	0.0000
2	1.0000
3	0.0000
4	1.0000
5	0.0000
6	1.0000

Covariate #2: mid_release
not Period Specific

Population	
1	0.0000
2	0.0000
3	1.0000
4	1.0000
5	0.0000
6	0.0000

This text window **(Data -> Data Listing)** displays the input data.

An example of a data listing for **chinook.dat** is displayed. As the data listing was too long to display in a single graphic, portions of the listing are shown in two **different** graphics. In the upper graphic, the data listing gives the name of the data file (/mounts/ **ruddy/surph/public/chinook.dat),** the number of populations in the data set **(6),** the number of sampling periods **(4),** the number and names of the group covariates, and the number and names of the individual covariates. Also displayed in the upper graphic are two examples of time independent group covariates. The **first** group covariate is treatment. Indicator variables denote which populations are (l's) and are not (O's) within the treatment group. Thus, populations 2, 4, and 6 are the treatment populations, whereas populations 1, 3, and 5 are non-treatment populations. Similarly, for the second group covariate (mid-release), populations 3 and 4 are mid-release groups, whereas populations 1, 2, 5, and 6 are not mid-release groups. Since both of these group covariates were time-independent, a single value is given for each population.

In the lower graphic, **ATPase** is a time-dependent group covariate. The values of **ATPase** must be defined for each period for'each population. For population 1, the mean **ATPase** changes from a value of 8.3 in **the** 1st period, to 25.5 in 2nd period, to 29.1 in 3rd period, and to 28.8 in 4th period. Below the **ATPase** values, capture histories and the values of the individual covariates are displayed. The **first** 15 animals in population 1 are displayed. Each line within this portion of the data listing **pertains** to an individual animal. For example, the 6th animal in **chinook.dat** has the unique ID code **7F7F112F77** and has an associated capture history of 11000. This animal **(7F7F112F77)** was 124 mm in length, weighed 21.3 grams, and had a condition index of 1.120 at first capture.

Data Listing

Print Quit

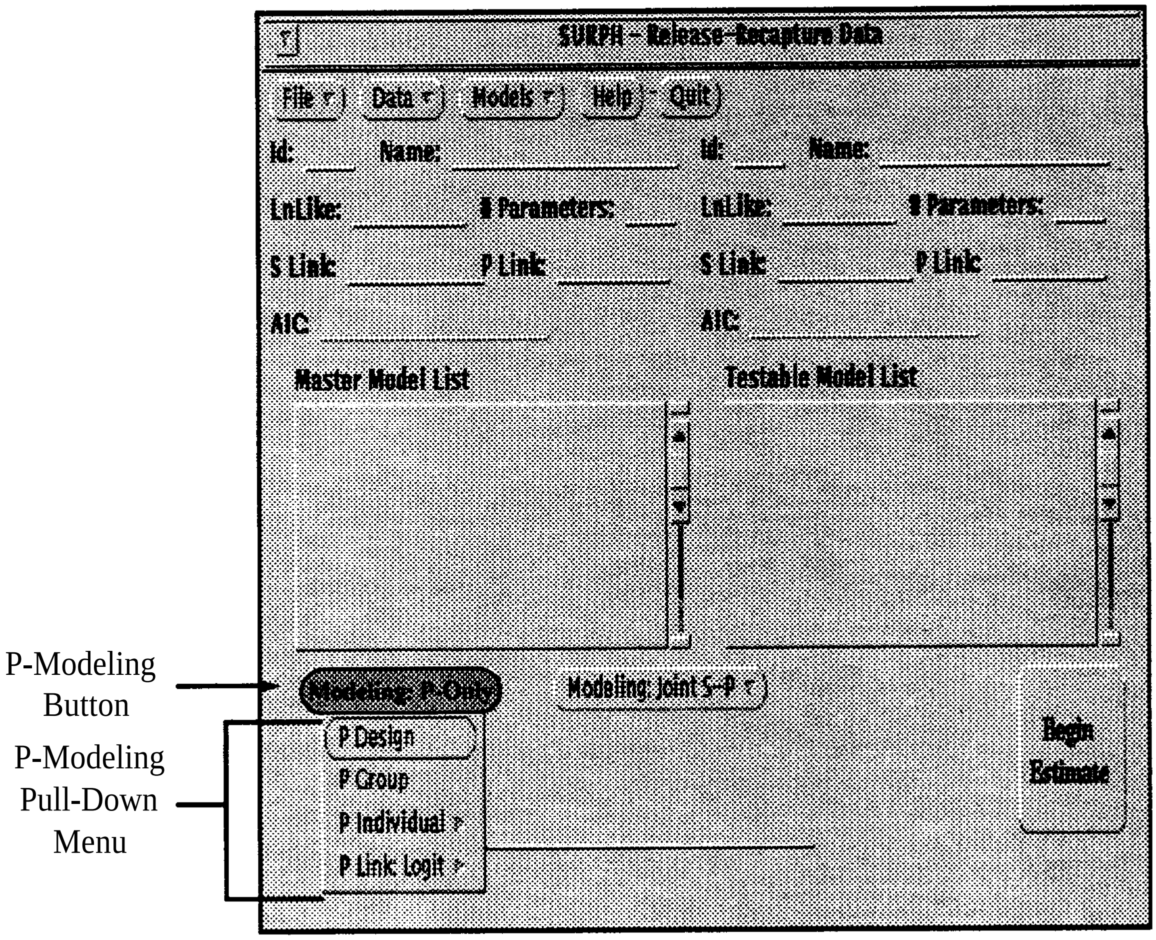
Covariate #4: ATPase

Population	Period 1	2	3	4
1	8.3000	25.5000	29.1000	28.8000
2	9.0000	20.4200	25.9000	33.0200
3	9.3000	22.5600	34.1400	39.2800
4	7.5000	20.2000	29.0000	23.6000
5	8.5000	21.4000	34.0000	25.3000
6	9.0000	21.3000	23.9000	23.6000

Capture Histories and Individual Covariates

Population 1

Animal	Id Code	Capture History	Individual Cuvariates: length	weight	condix
1	7F7E644A5A	10000	141.000	31.400	1.120
2	7F7F0F554E	10000	144.000	33.600	1.120
3	7F7F0F5C7D	11000	156.000	39.900	1.040
4	7F7F103974	10100	146.060	34.900	1.110
5	7F7F112C28	10100		29.500	1.090
6	7F7F112F77	11000	128.000	21.300	1.120
7	7F7F116C67	10100	144.000	29.400	0.960
8	7F7F117B41	10010	127.000	23.309	1.150
9	7F7F12130C	10060	134.000	27.200	1.120
10	7F7F123C48	10001	124.000	22.900	1.210
11	7F7F124776	10000	93.000	9.100	1.120
12	7F7F126E1C	10100	115.000	16.600	1.040
13	7F7F141B11	11100	127.600	24.200	1.190
14	7F7F141B39	11100	146.000	31.600	1.020
15	7F7F141C2D	10000	133.000	25.700	1.090



This is the pull-down menu that is displayed when you right-click on the **P-Modeling Button.** Using this menu, you create models of capture probabilities independent of the survival process (see Section 4.4.1 in the SURPH.l manual).

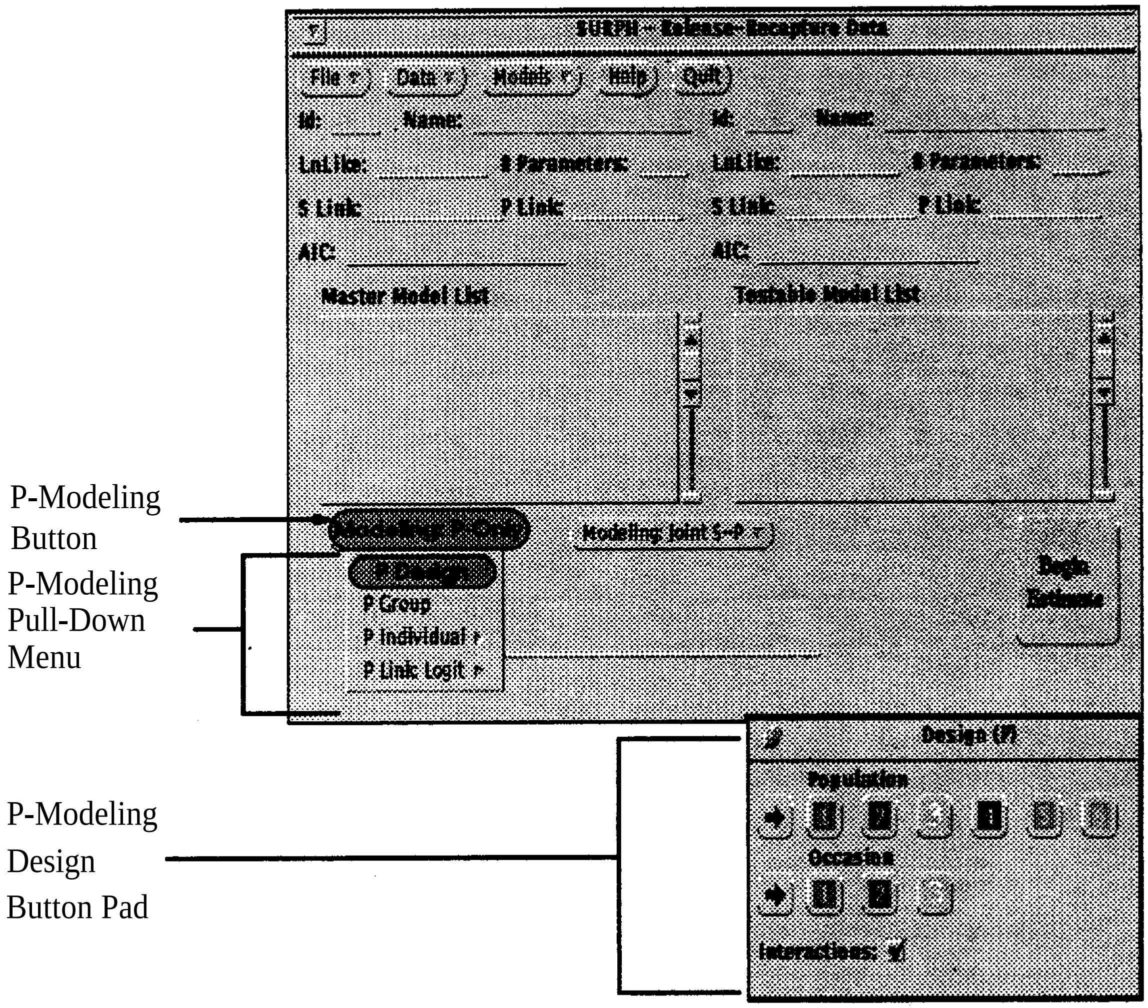
The menu choices available **from** this pull-down menu specify elements of models of capture probabilities:

P Design - Specify effects of experimental design factors. A Button Pad will appear that allows you to specify period and population-specific capture parameters **(see** Capture Modeling: Design Factors **).**

P Group Specify group covariate effects. A Button Pad will appear that allows you to specify parameters for period-specific group covariate effects on capture probabilities (see Capture Modeling: Group Covariates).

P Individual - **Specify** individual covariate effects. For **each** covariate selected, a Button Pad will appear. This Button Pad allows you to specify parameters for period and **population-specific** effects on capture probabilities (see Capture Modeling: Individual Covariates).

P Link - Specify the function that links covariate effects **to** capture probabilities. The default link is **Logit.** Alternatively, the Hazard link function can be specified (see Capture Modeling: Link Function).

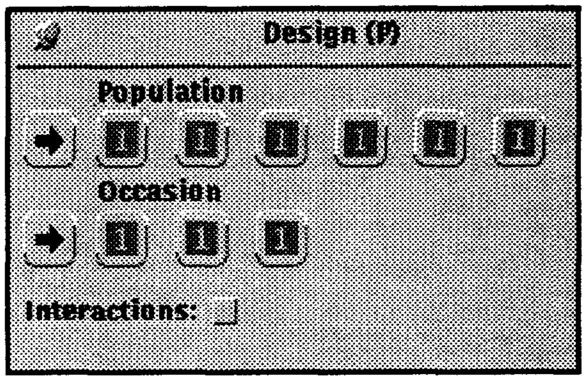


The P-Design Button Pad **(Modeling: P-Only -> P Design** or **Modeling: Joint S-P -> P Design)** allows you to model capture probabilities as a function of the design factors population and occasion. Unique button settings within the population or occasion row indicate unique effects for each population or occasion, whereas duplicated button settings within the population or occasion row indicate a shared parameter. For example, in the Design Button Pad shown above, all occasion and population effects are estimated with unique parameters, for there are unique settings on each button within a row, and there are interactions among the populations and occasion effects. Left-click on the individual buttons or use the Quick Buttons (see Quick Buttons - Instructions) to change button settings.

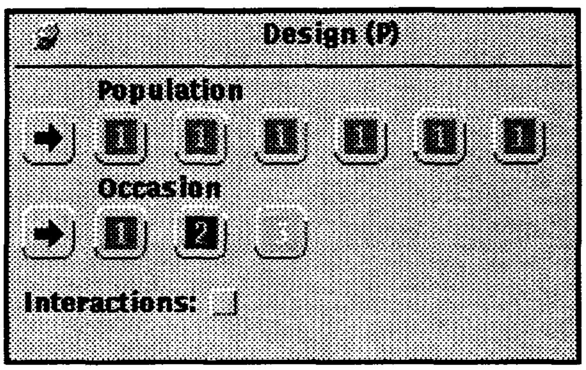
The example above shows **Model Pst, with main effects** for space (population) and time (occasion) **and** interactions among all population and occasion effects. Other standard capture models are illustrated under

Capture Modeling - Examples

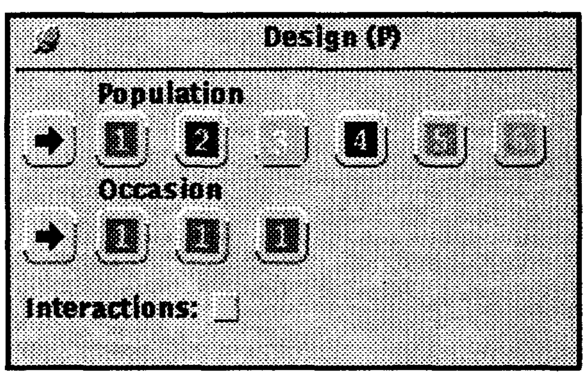
Example of Four Major Design Factor Models



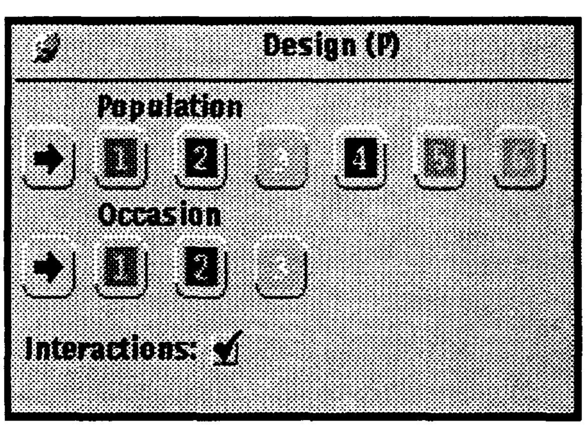
This display illustrates the simplest of the capture models. Only one parameter is **estimated;** a common capture probability for all populations and on all occasions. This is also known as Model **P.**



This display illustrates **Model Pt**. Under this model, there is a common capture parameter across populations on each sampling occasion, but capture probabilities vary across occasions.

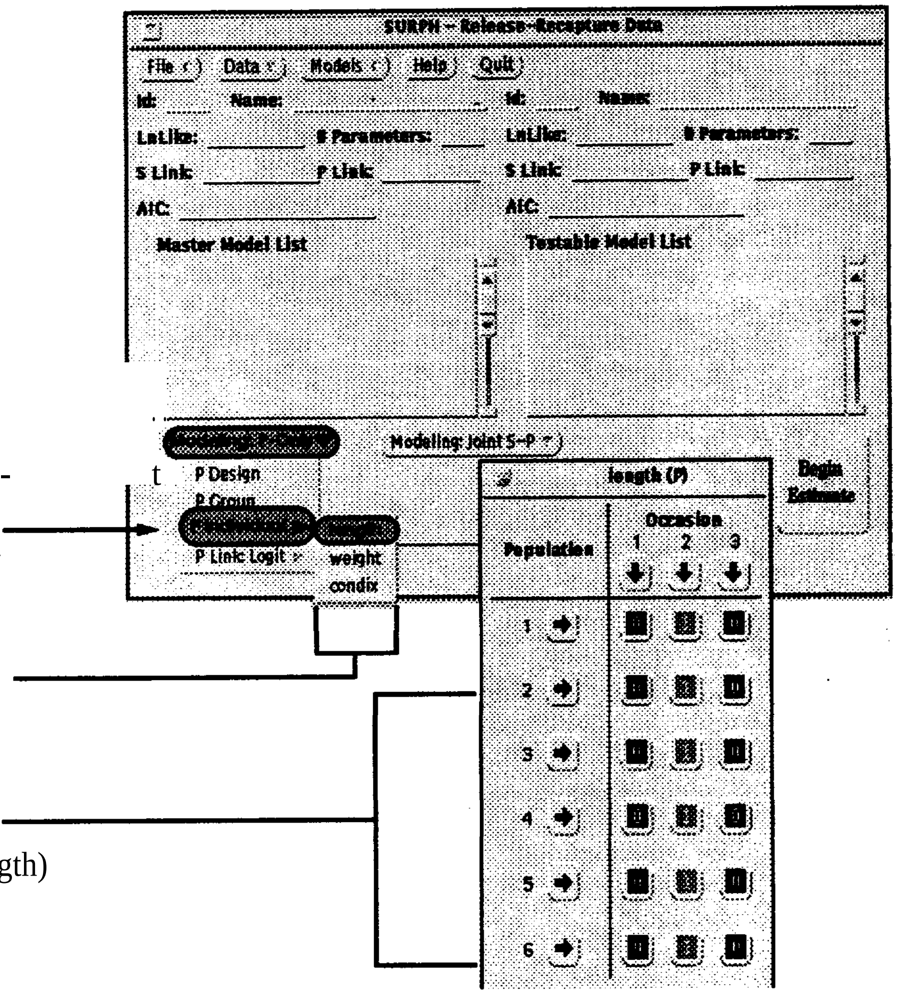


This display illustrates **Model Ps.** Under this model, the capture probability for each population is constant over all occasions, but capture probabilities vary among populations.



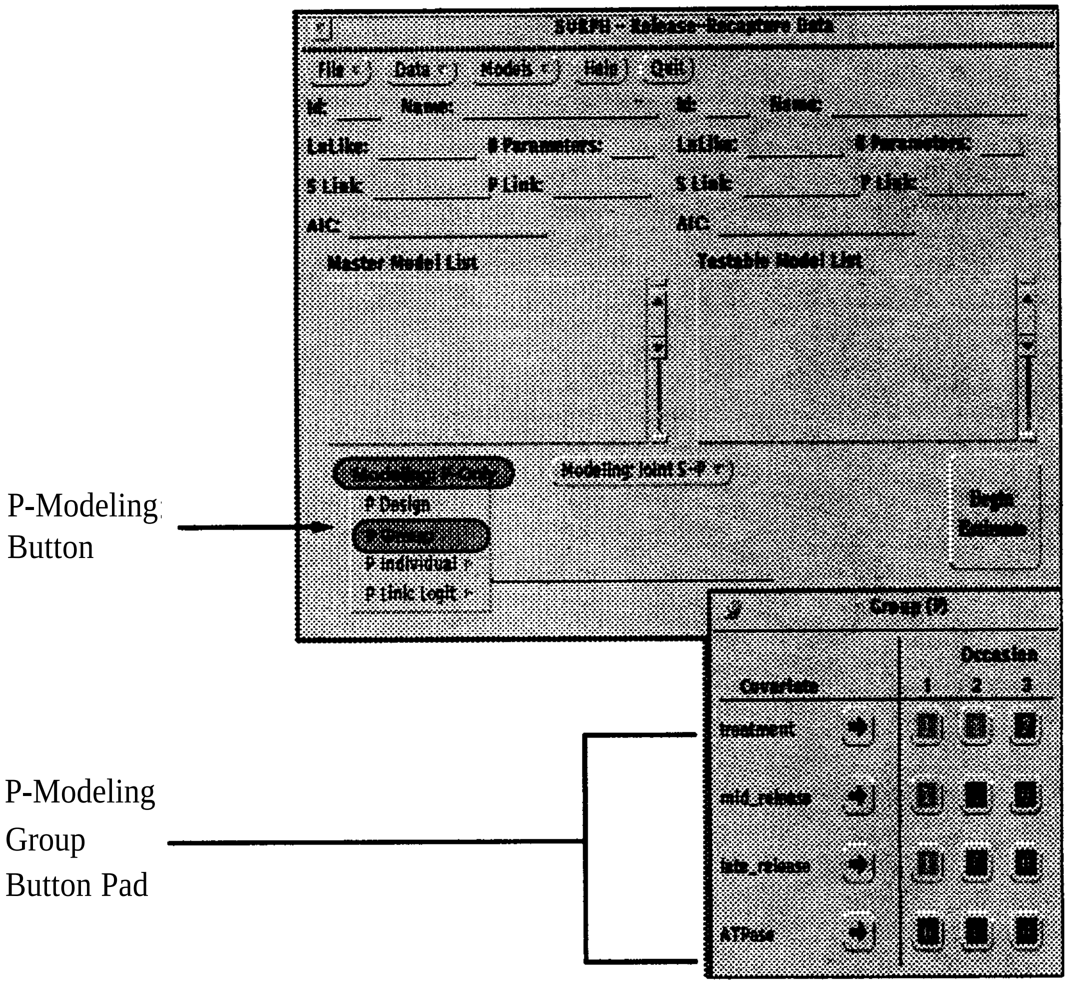
The final display illustrates the most general of design factor models. Under **Model Pst,** capture probabilities are unique on each occasion for each population. Note that the interaction **checkbox** must be checked to fully **parameterize** this option. If interactions are not checked, the population and occasion effects are additive (a “main effects” model).

THIS PAGE INTENTIONALLY LEFT BLANK



The P-Modeling Individual Covariate Button Pad **(Modeling: P-Only -> P Individual** or **Modeling: Joint S-P -> P Individual) allows you** to model capture probabilities as a function of Individual Covariates. You access the **P-Modeling Individual Button Pad** by selecting the covariate of interest from the submenu under **P Individual. You** change the **parameterization** of the capture model by altering the button settings within the **P-Modeling Individual Button Pad.** A zero indicates that no effect is estimated for the covariate on that occasion for that population, whereas non-zero settings indicate that a parameter is estimated for the effect of the covariate. If the same button setting is used for a particular covariate, a common parameter is estimated for the effect of the covariate for all populations and on all occasions sharing the setting. If different button settings are used, separate parameters are estimated. **Left-click** on the individual buttons or use the Quick Buttons

(see Quick Buttons - Instructions **) to change button settings.**



The Group Covariate Button Pad **(Modeling: P-Only -> P Group** or **Modeling: Joint S-P -> P Group)** allows you to model capture probabilities as a function of Group Covariates. You change the **parameterization** of the capture model by altering the button settings on the Button Pad. A zero indicates that no effect is estimated for that covariate on that **occasion,** whereas non-zero settings indicate that a parameter is estimated for the effect of that covariate. If the same button setting is used for more than one **occasion** for a particular covariate, a common parameter is estimated for the effect of the covariate on all occasions sharing the setting. If different button settings are used, separate parameters are estimated. Left-click on the individual buttons or use the Quick Buttons (see Quick Buttons - Instructions **) to change button settings.**

The model specified above has a common treatment effect on occasions 1 and 2, and a different treatment effect on occasion 3. In addition, there is a mid-release effect on occasion 1, but not on occasions 2 or 3, and a late-release effect on occasion 1, but not on occasions 2 or 3. Note that the treatment, mid_release and late-release settings are independent. The "1" settings for treatment on occasions 1 and 2, the "1" setting for mid-release on occasion 1, and the "1" setting for late-release on occasion 1 do not indicate that the same parameter is used for the three covariates.

The current example shows a capture made1 with a common length covariate effect for all populations on **occasion** 2 (all “l’s”), and no length effects on other occasions (all “O’s”).

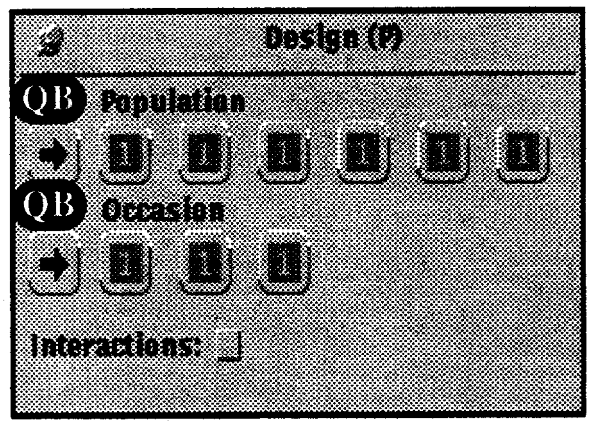
Illustration of Quick Button Use for Design Factors

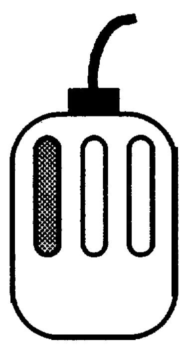
The quick buttons (arrow keys **labelled** QB in the diagram) allow you to change the parameters in the model (button settings) quickly. The use of quick buttons (QB) for Design Factor Button Pads is illustrated below. The use of quick buttons for all other Button Pads is illustrated on the following page.

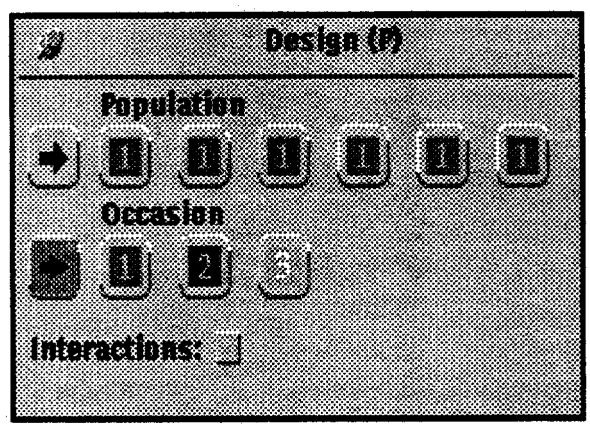
Left-clicking on the Quick Button alternates between a single parameter for all buttons in the row, and a distinct parameter for each button.

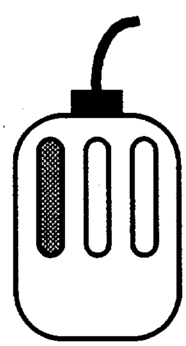
Hence, to go from the first illustrated setting to the second, left-click on the QB for Occasions. Note, the display changes from "111" to " 123".

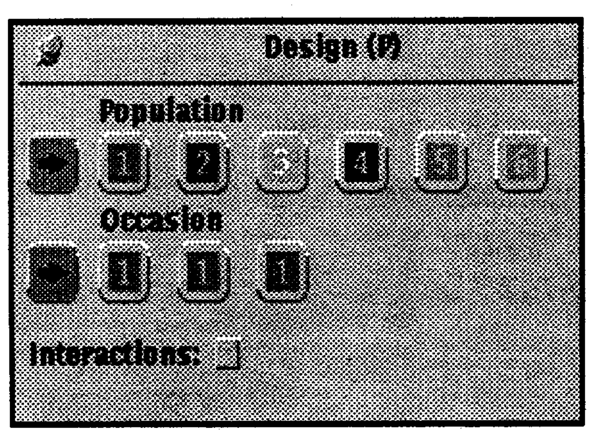
Similarly, to go from the second to the third illustrated setting, left-click on the QB for Populations, and then left-click on the QB for Occasions. This changes the Population row from "1111" to **"123456",** and changes the Occasion row from "123" to "111".









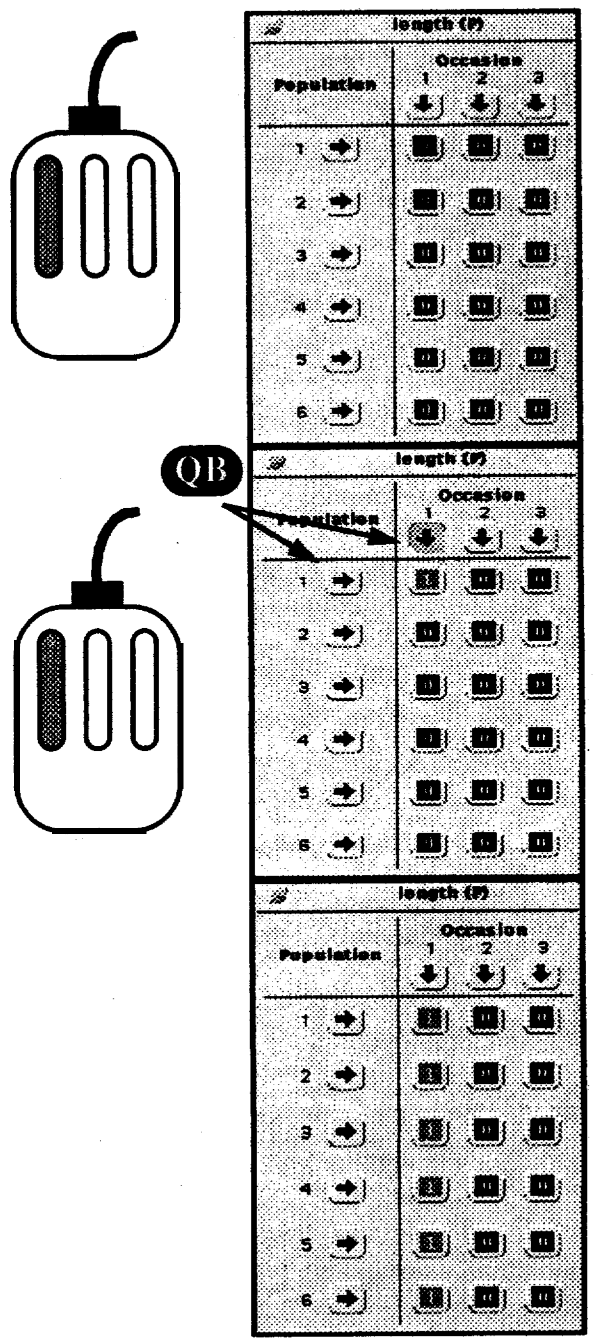


THIS PAGE INTENTIONALLY LEFT BLANK

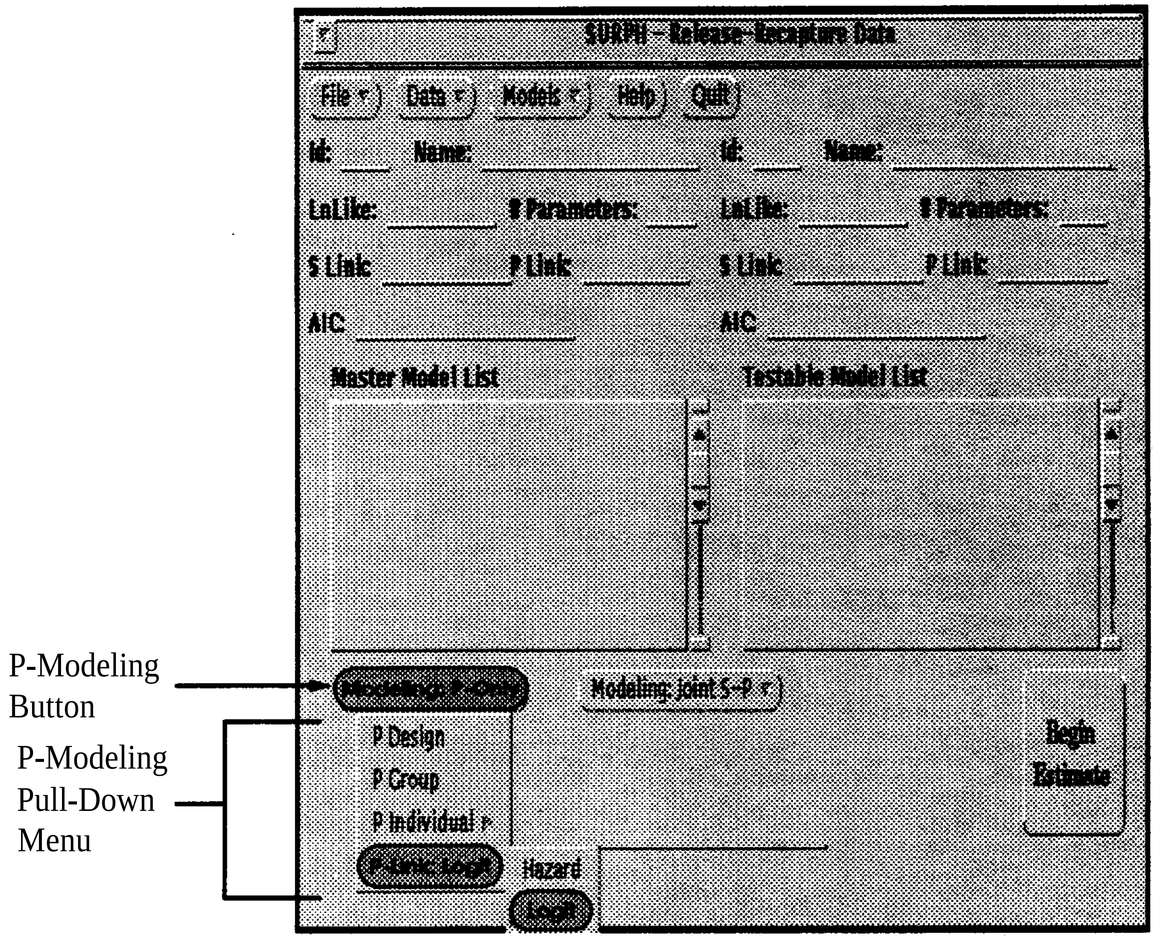
Illustrations of Quick Button Use for Group and Individual Covariate Modeling

The quick buttons (QB) for the group and individual covariate modeling respond slightly differently than the quick buttons for the Design Factors (see previous page). The quick button for individual covariates copies the parameter designation on the first button in the row or column to all buttons within that row or column.

In the example, the first model has no individual parameters fit in the capture model (all buttons set to “0”). Left-clicking on the button in the first row-first column changes that button from a **“0”** to a “1”. Left-clicking on the “Occasion” quick button in column 1 copies the “1” in the first row-first column throughout the entire first column. You can copy a setting across rows similarly.

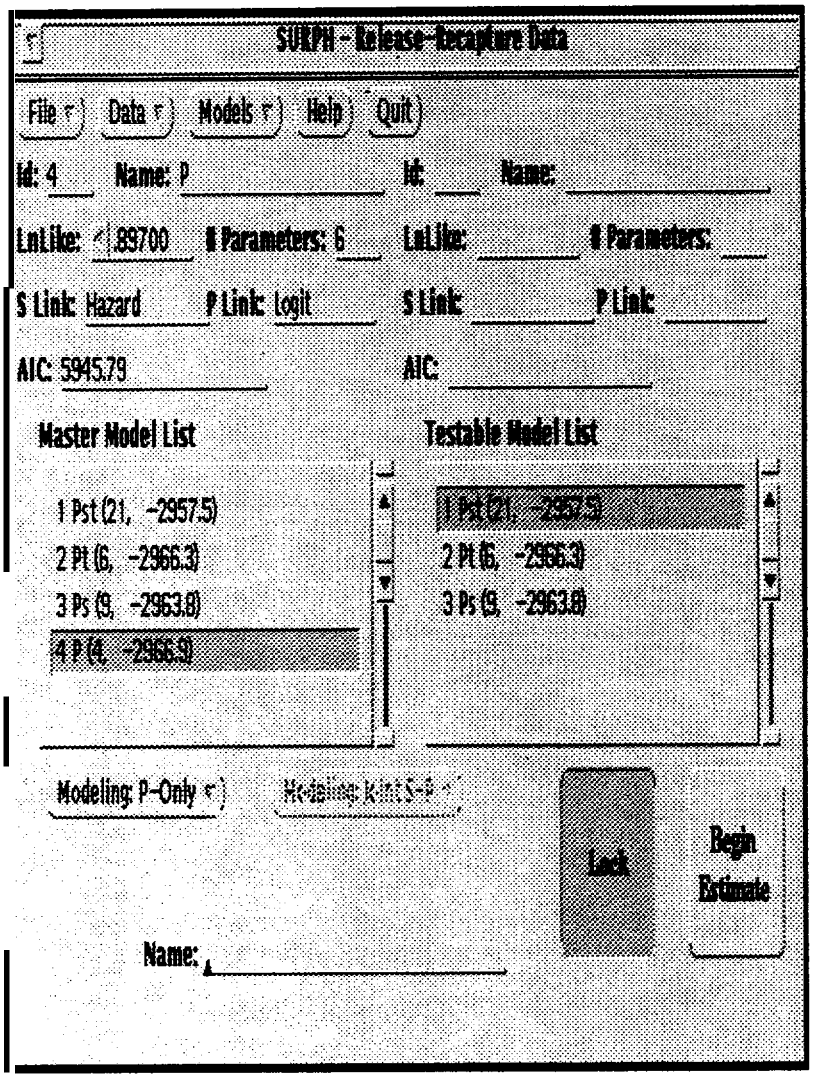


THIS PAGE INTENTIONALLY LEFT BLANK



The P-Modeling **Link** Function **(Modeling: P-Only -> P-Link** or **Modeling: Joint S-P -> P-Link) allows you** to specify whether effects on the capture will be linked to capture probabilities using a Hazard Function link or a **Logit** Function link The default **link** for capture modeling is **Logit**.

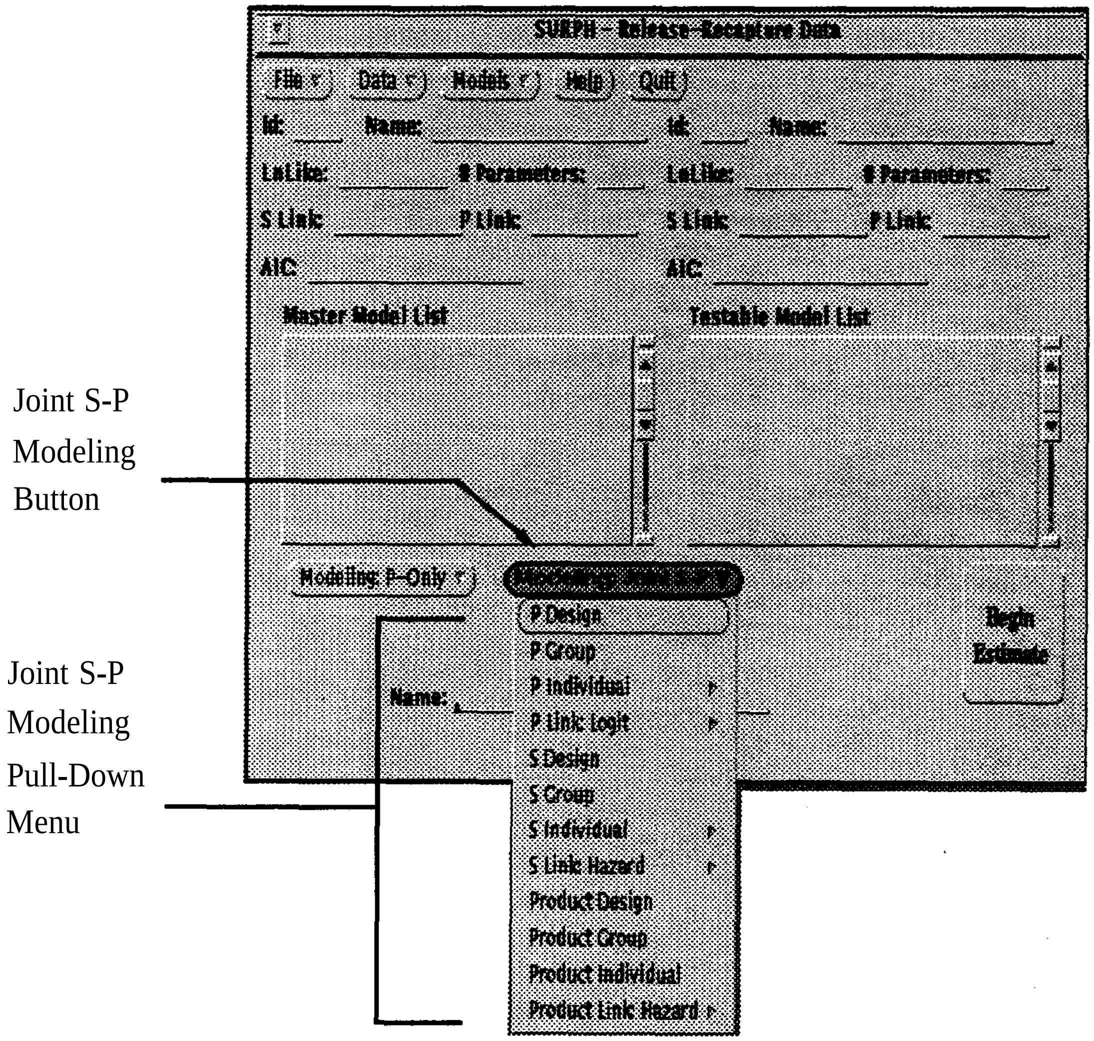
THIS PAGE INTENTIONALLY LEFT BLANK



This graphic shows the typical appearance of the SURPH Base Window after you estimate a series of models for capture probabilities **under Modeling: P-Only. After** a series of capture models is estimated, you have the opportunity to choose which capture model will be retained for the survival modeling portion of the data analysis. The capture model that is chosen is then "locked". Once the capture model has been locked (left-click on **the Lock Button),** you cannot change the model for capture probabilities during subsequent survival analysis.

In the above example, capture model **"P"** is selected in the Master Model List. Left-click on the **Lock Button** to specify the "P" model for capture probabilities. All subsequent models of survival specified under the Modeling: Joint S-P Button will have the "P" model specified for capture probabilities.

THIS PAGE INTENTIONALLY LEFT BLANK



This is the pull-down menu displayed when you right-click on the **Modeling: Joint S-P Button.** Using this menu, you create models of capture and survival jointly (see Section 4.4.3 in the SURPH.l manual).

The commands available from this pull-down menu specify models for capture and survival probabilities:

P Design - Specify effects of the experimental design factors on capture probabilities. A Button Pad will appear that allows you to specify occasion and population-specific effects (see Capture Modeling, Design Factors).

P Group Specify group covariate effects on capture probabilities. A Button Pad will appear that allows you to specify **parameters** for **occasion-specific** group covariate effects
(see Capture Modeling: Group Covariates **).**

P Individual - Specify individual covariate effects on capture probabilities. For each covariate selected, a Button Pad will appear. This Button Pad allows you to specify parameters for occasion and **population-specific** effects
(see Capture Modeling: Individual Covariates **).**

P Link Specify the function that links covariate effects to capture probabilities. The default link is **Logit.** Alternatively, the Hazard link function can be specified
(see Capture Modeling: Link Function **).**

S Design Specify effects of the experimental design factors on survival probabilities. A Button Pad will appear that allows you to specify period and population-specific effects on survival probabilities
(see Survival Modeling: Design Factors **).**

S Group Specify group covariate effects on survival probabilities. A Button Pad will appear that allows you to specify parameters for period-specific group covariate effects
(see Survival Modeling: Group Covariates **).**

S Individual - Specify individual covariate effects on survival probabilities. For each covariate selected, a Button Pad will appear. This Button Pad allows you to specify parameters for period and

population-specific effects
(see Survival Modeling: Individual Covariates **).**

S Link - Specify the function that links **covariate effects** to survival probabilities. The default **link** is Hazard. Alternatively, the **Logit** link function can be specified
(see Survival Modeling: Link Function **).**

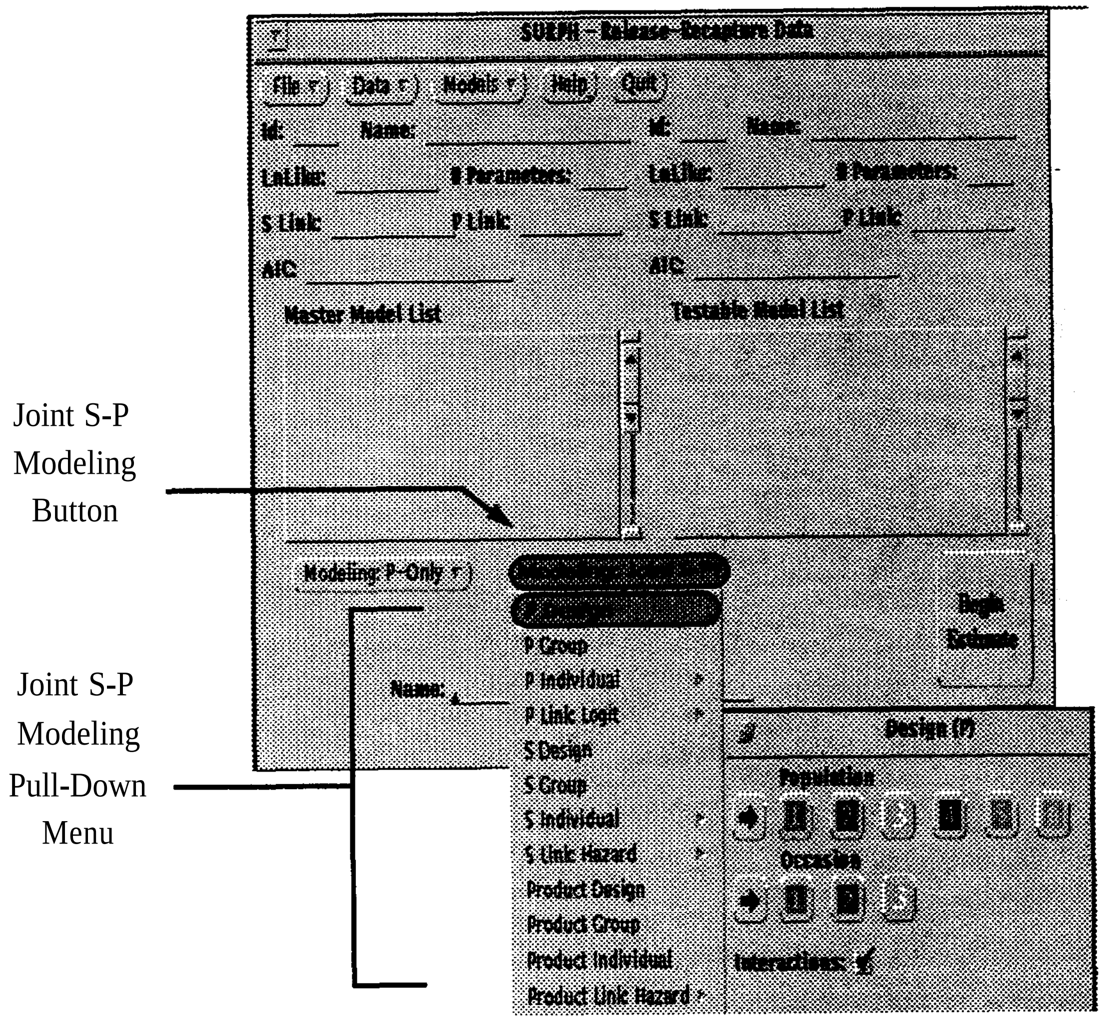
Product Design - Specify effects of the experimental design factors on the final period capture-survival probabilities. A Button Pad **will** appear that allows you **to** specify population effects
(see Product Modeling: Design Factors **).**

Product Group - Specify group covariate effects on the final period capture-survival probabilities. A Button Pad will appear that allows you to specify parameters for group covariate effects
(see Product Modeling: Group Covariates **).**

Product Individual - Specify individual **covariate** effects on the final period capture-survival probabilities. A single Button Pad will appear that includes all individual **covariates.** This Button Pad allows you to specify population-specific effects
(see Survival Modeling: Individual Covariates **).**

Product Link - **Specify** the function that links covariate effects to the final period capture-survival probabilities. The default **link** is Hazard. Alternatively, the **Logit** link function can be **specified**
(see Product Modeling: Link Function **).**

THIS PAGE INTENTIONALLY LEFT- BLANK



This Button Pad **(Modeling: Joint S-P -> P Design) allows you to** model capture probabilities as a function of the experimental design factors for Population and Occasion. This function is fully described **under** Capture Modeling: Design Factors

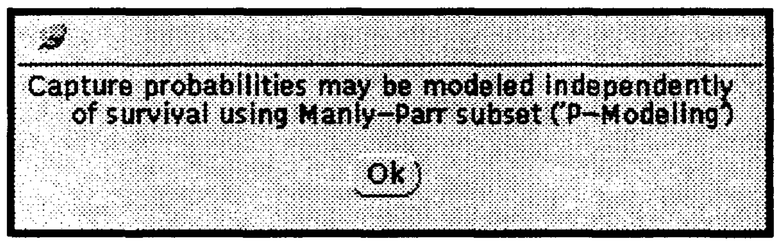
The current example shows **Model Pst, with** main effects for space (population) and time (occasion), and interactions among all population and occasion effects. Other standard capture models are illustrated under

Capture Modeling - Examples .

There is a subtle difference between modeling the capture process using the **Modeling: Joint S-P Button,** rather **than** **the Modeling: P-Only Button.** Under the **Modeling: P-Only Button, an** extension of the Manly-Parr (1968) model is used to model the **capture** process. Only a subset of the data is used (see Data Summary: Manly-Parr).

Under P-Only modeling, the models for the capture probabilities are independent of the survival process, and no survival model is implied during the capture modeling (see Section **4.4.1 in the** SURPH. 1 manual). Under **the Modeling: Joint S-P Button,** however, capture and survival are **modelled** jointly. Thus the capture probabilities are not independent of the survival process.

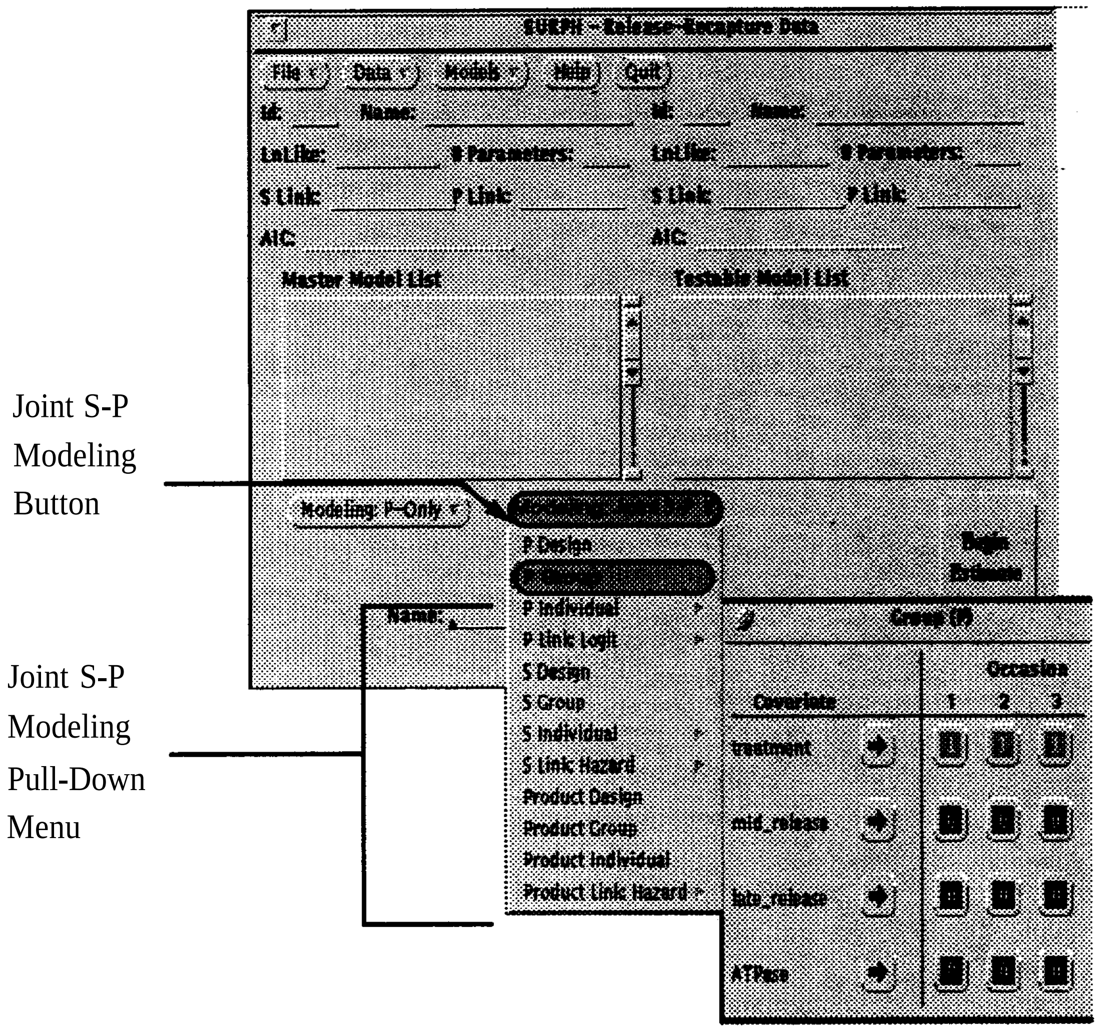
Conclusions regarding the capture model might depend on the jointly-defined survival model. Simulation studies have shown that if **Model Sst is used** for survival probabilities (unique parameters for all period and population main effects and interactions) conclusions regarding the capture model will generally be identical whether **P-Only** or **Joint S-P** modeling is used.



This pop-up message informs you that there are two separate ways to approach modeling the capture histories (see Joint Modeling: Capture Design Factors **).**

The pop-up message appears when you select a capture modeling option **from** the **Modeling: Joint S-P** menu before selecting any survival modeling option. The message is a reminder that the capture probabilities can be modeled independently.

THIS PAGE INTENTIONALLY LEFT BLANK



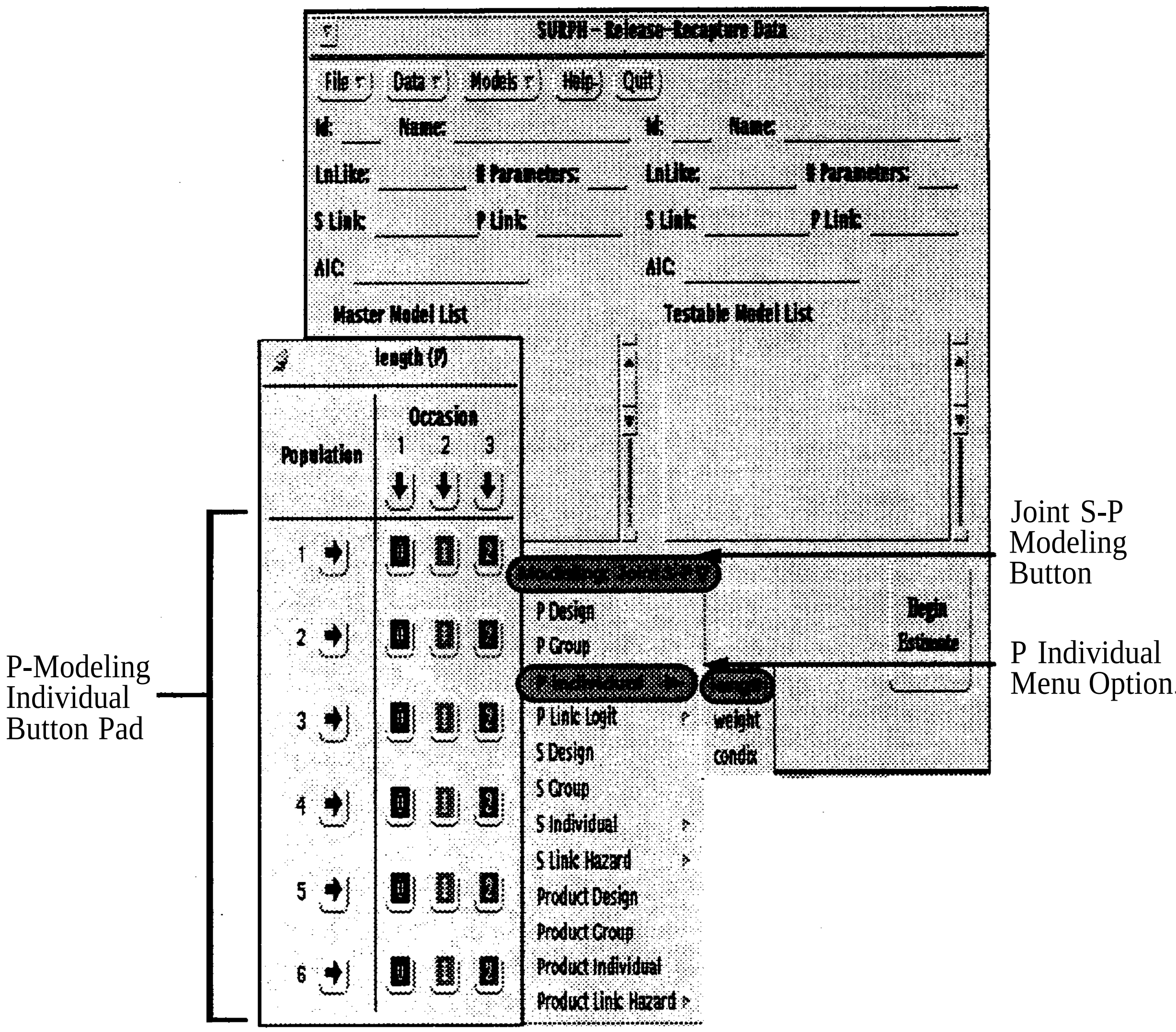
The P-Modeling Group Covariate Button Pad **(Modeling: Joint S-P -> P Group** or **Modeling: P-Only -> P Group)** allows you to model capture probabilities as a function of Group Covariates. This function is fully described under

Capture Modeling: Group Covariates .

The current example shows a capture model with a common treatment covariate effect throughout all occasions (all "l's" in the upper row), and no other group covariate effects (all "O's" elsewhere).

4.86

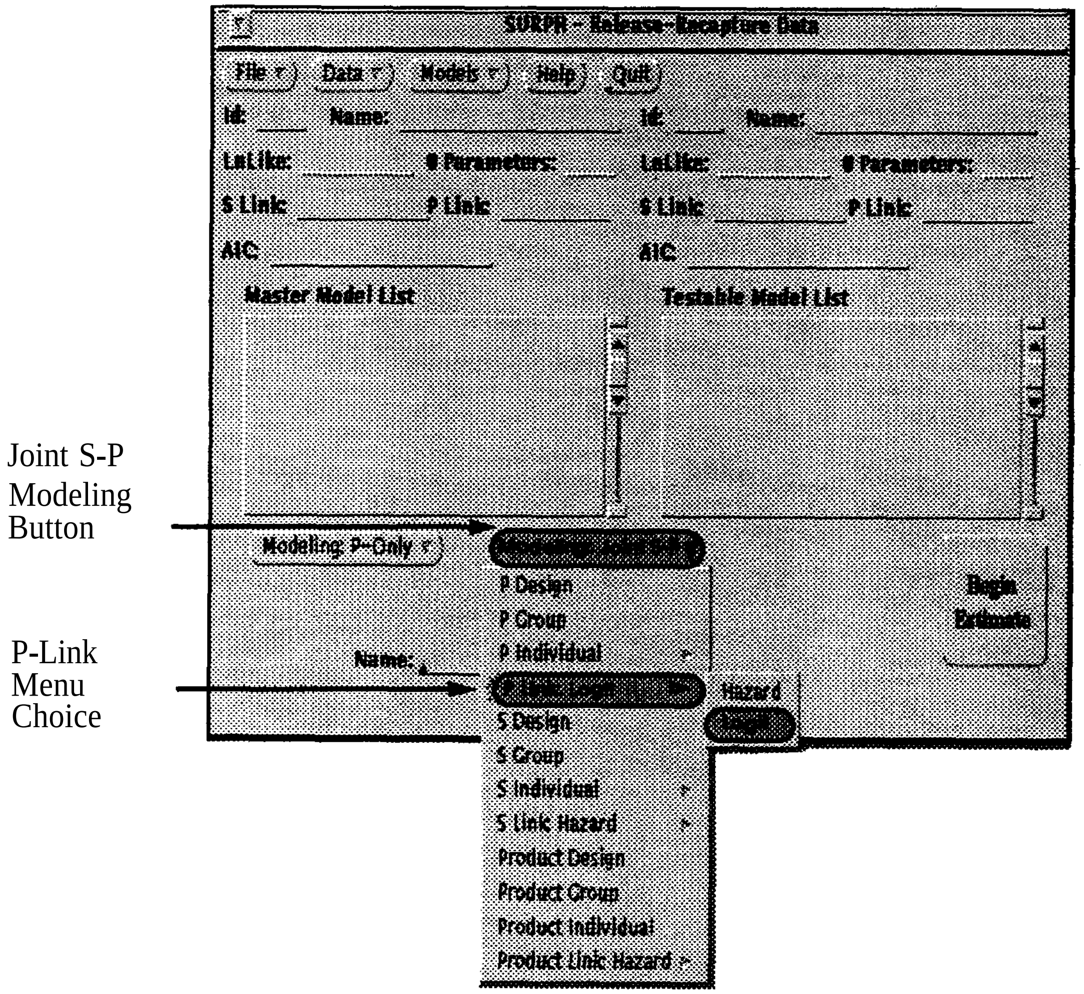
THIS PAGE INTENTIONALLY LEFT BLANK



The P-Modeling Individual Covariate Button Pad **(Modeling: Joint S-P -> P Individual** or **Modeling: P-Only -> P Individual) allows you to** model capture probabilities as a function of Individual Covariates. This function is fully described under Capture Modeling: Individual Covariates .

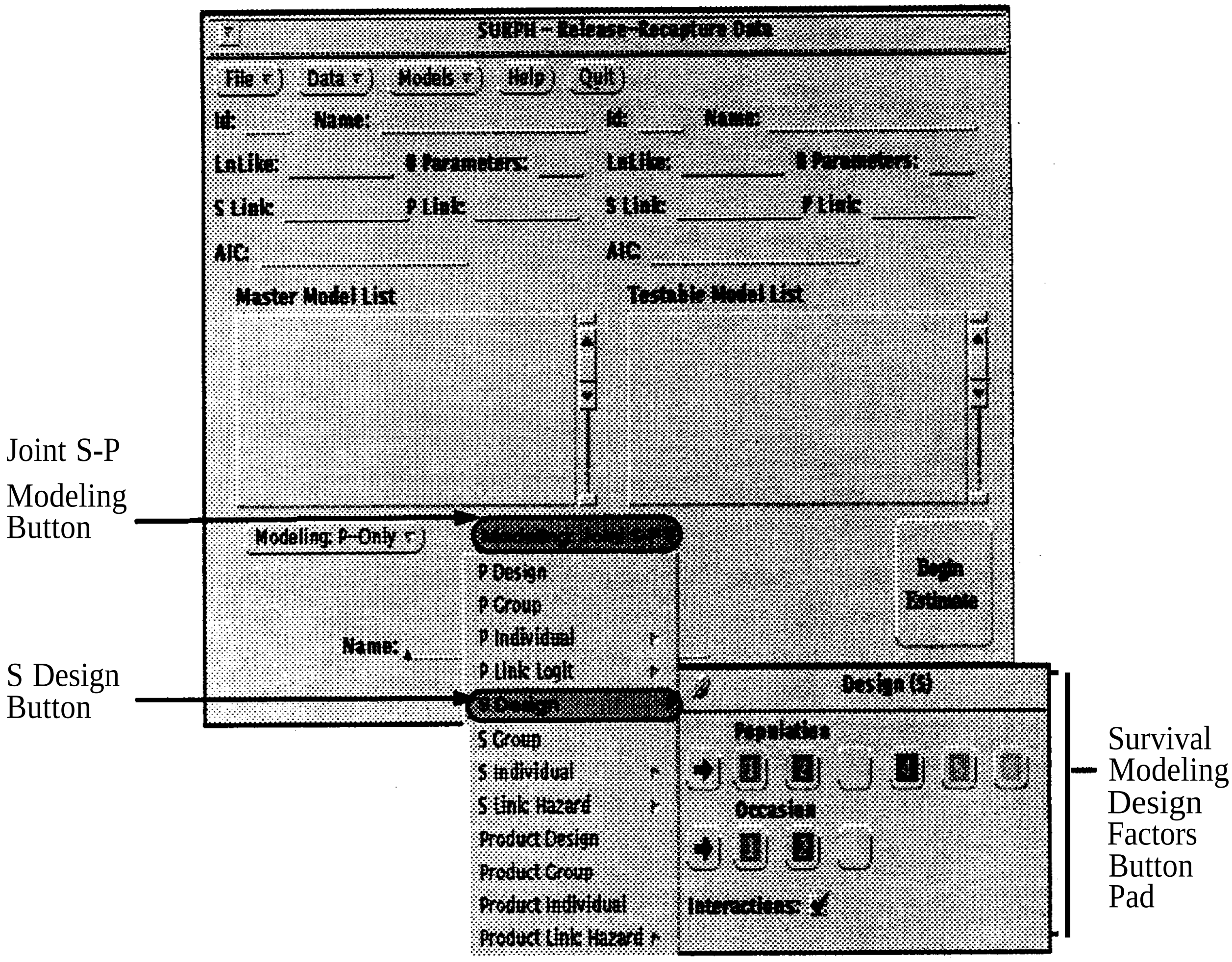
The current example shows a capture model with no **length** effects on occasion 1 (all "O's" in the first column), a common **length** effect for all populations on occasion 2 (all "1 's" in the second column), and a second common, but different, length effect on occasion 3 (all "2's in the third column).

THIS PAGE INTENTIONALLY LEFT BLANK



The P-Modeling Link **Function (Modeling: Joint S-P -> P-Link or Modeling: P-Only -> P-link) allows you** to specify whether effects on the capture will be linked to capture probabilities using a Hazard Function link or a **Logit** Function link. The default link for capture modeling is **Logit**.

THIS PAGE INTENTIONALLY LEFT BLANK

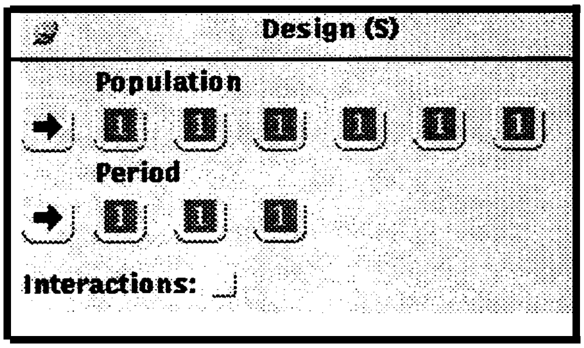


The S-Design Button Pad **(Modeling: Joint S-P -> S Design) allows you** to model survival probabilities as a function of the design factors population and period. Unique button settings within the population or period row indicate unique effects for each population or period, respectively, whereas duplicated button settings within the population or period row indicate shared parameters. **In** the example above, all population and period main effects and interactions are estimated with unique parameters, for there are unique button settings on each button within a row, and the interaction box has been checked. Left-click on the individual buttons or use the Quick Buttons (see Quick Buttons - Instructions **) to change button settings.**

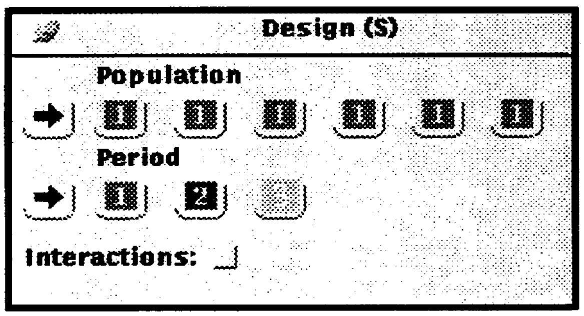
The example above shows **Model Sst** for survival probabilities, with main effects for space (population) and time (period) and interactions among all population and period effects. Other standard survival models are illustrated under Survival Modeling - Examples .

THIS PAGE INTENTIONALLY LEFT BLANK

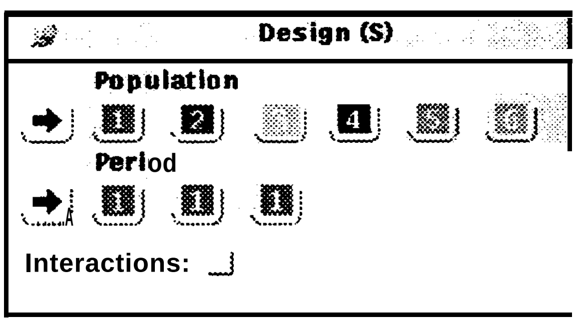
Example of Four Major Design Factor Models



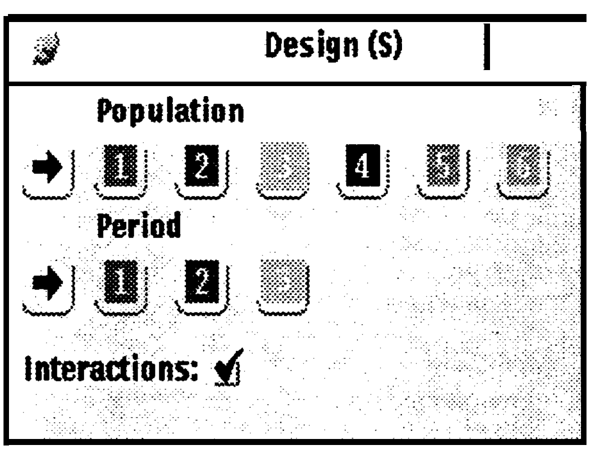
This display illustrates the **simplest** of the survival models. Only one parameter is estimated; a common survival probability for all populations and in all **periods. This is also known as Model S.**



This display illustrates **Model St.** Under this model, there is a common survival parameter across populations during each sampling period, but survival probabilities differ among periods.

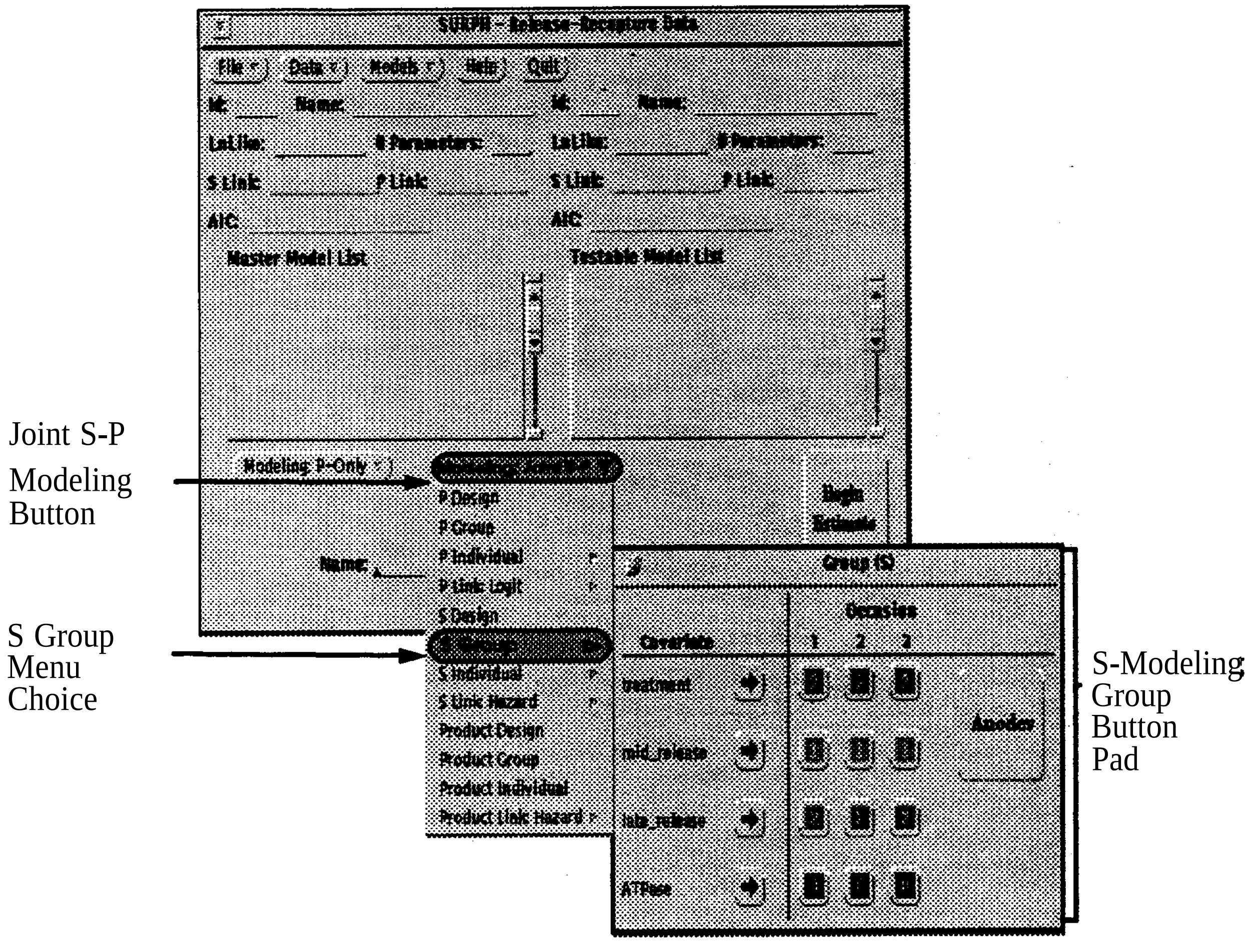


This display illustrates **Model Ss.** Under this model, the survival probability for each population is constant across all periods, but survival probabilities vary among populations.



The final display illustrates the most general of design factor models. Under **Model** Sat, survival probabilities are unique during each period for each population. Note that the interaction **checkbox** must be checked to fully **parameterize** this option. If interactions are not checked, the population and period effects **are** additive (a “main effects” model).

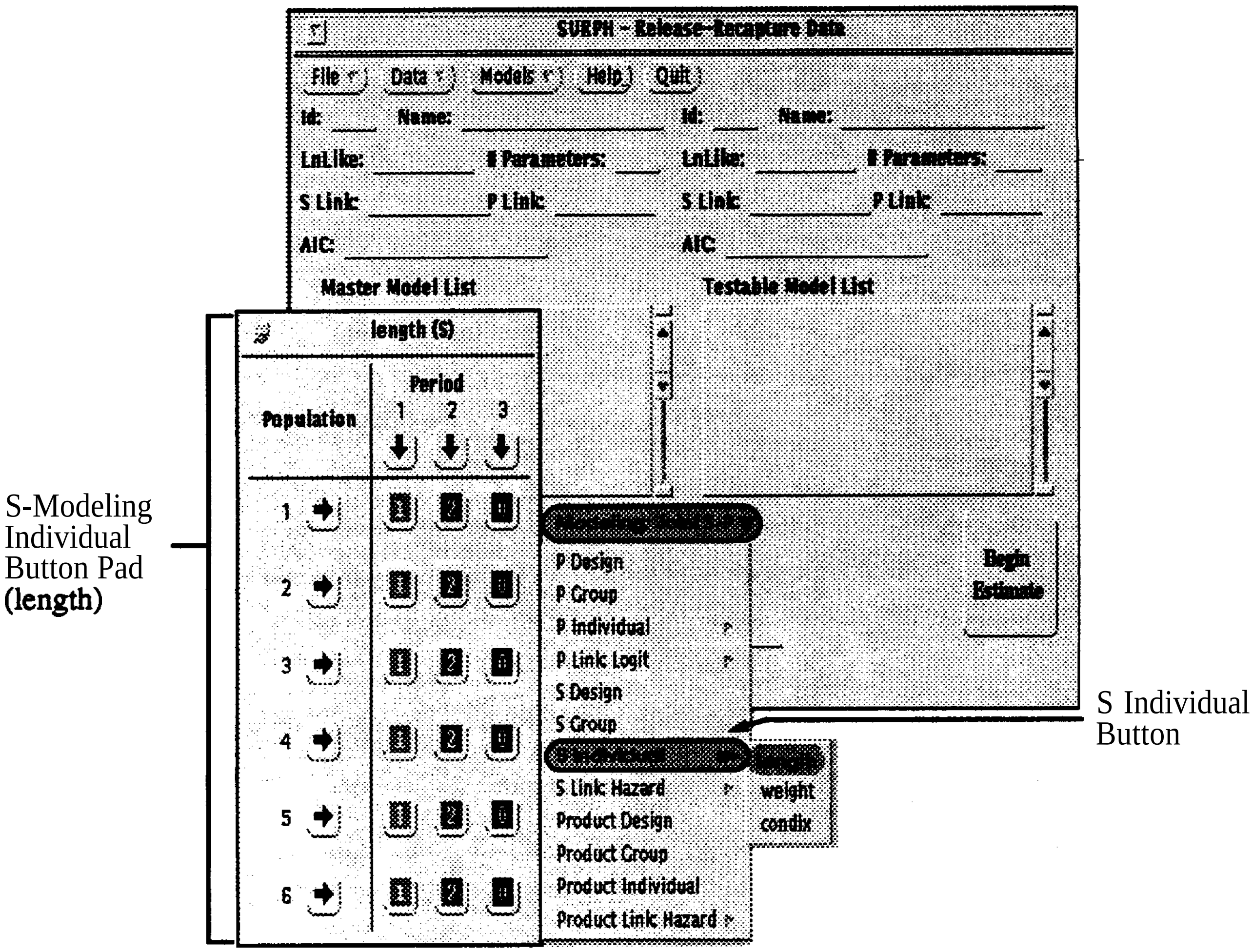
THIS PAGE INTENTIONALLY LEFT BLANK



The Survival Modeling Group Covariate Button Pad **(Modeling: Joint S-P -> S Group)** allows you to model survival probabilities as a function of Group Covariates. You change the **parameterization** of the survival model by'altering the button settings on **the S-Modeling Group Covariate Button Pad.** A zero indicates that no effect is estimated for that covariate in that period, whereas non-zero settings indicate that a parameter is estimated for the effect of that covariate. **If** the same button setting is used for more than one period for a particular covariate, a **common** parameter is estimated for the effect of the wvariate during all periods sharing the setting. **If** different button settings are **used, separate parameters are estimated. Left-click on the individual buttons or use the Quick Buttons (see** Quick Buttons - Instructions **) to change button settings.**

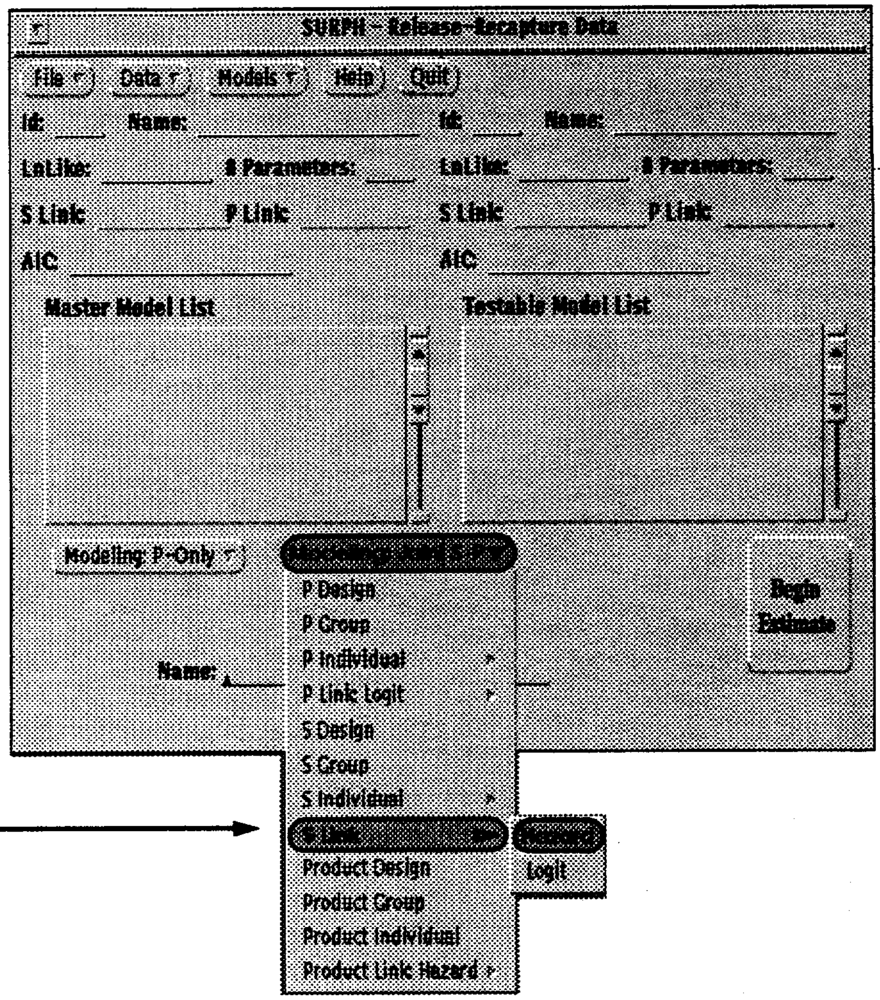
The model specified above has a common treatment effect during all periods (all “2’s” across the treatment row), a common mid-release effect during all periods (all “l’s” across the mid-release row), and a common late-release effect during all periods (ah “2’s” across the late-release row). Note that the treatment, mid-release and late-release settings are independent. The “2” settings for treatment and the **“2”** setting for late-release do not indicate that the same parameter is **used** for both **covariates.**

Alternatively, you can investigate the effects of Group Covariates using Analysis of Deviance. To use Analysis of Deviance to assess Group Covariate effects, le **..ck** on **the Anodev Button.** For further information on Analysis of Deviance, see ANODEV **and all pages associated with** Analysis of Deviance .



The Survival Modeling Individual Covariate -Button Pad **(Modeling: Joint S-P -> S Individual)** allows you to model survival probabilities as a function of Individual Covariates. You **access the S-Modeling Individual Button Pad** by selecting the wvariate of interest from the pull-right menu **under S Individual. You** change the **parameterization** of the survival model by altering the button settings within the **S-Modeling Individual Button Pad.** A zero indicates that no parameter is estimated for the wvariate in that period for that population, whereas non-zero button settings indicate that a parameter is estimated for the wvariate. If the same button setting is used for a particular covariate, a wmmon parameter is estimated for all populations and periods sharing that designation. If different button settings are used, separate parameters are estimated. L&t-click on the individual buttons or use the Quick Buttons to change button **settings (see** Quick Buttons - Instructions **).**

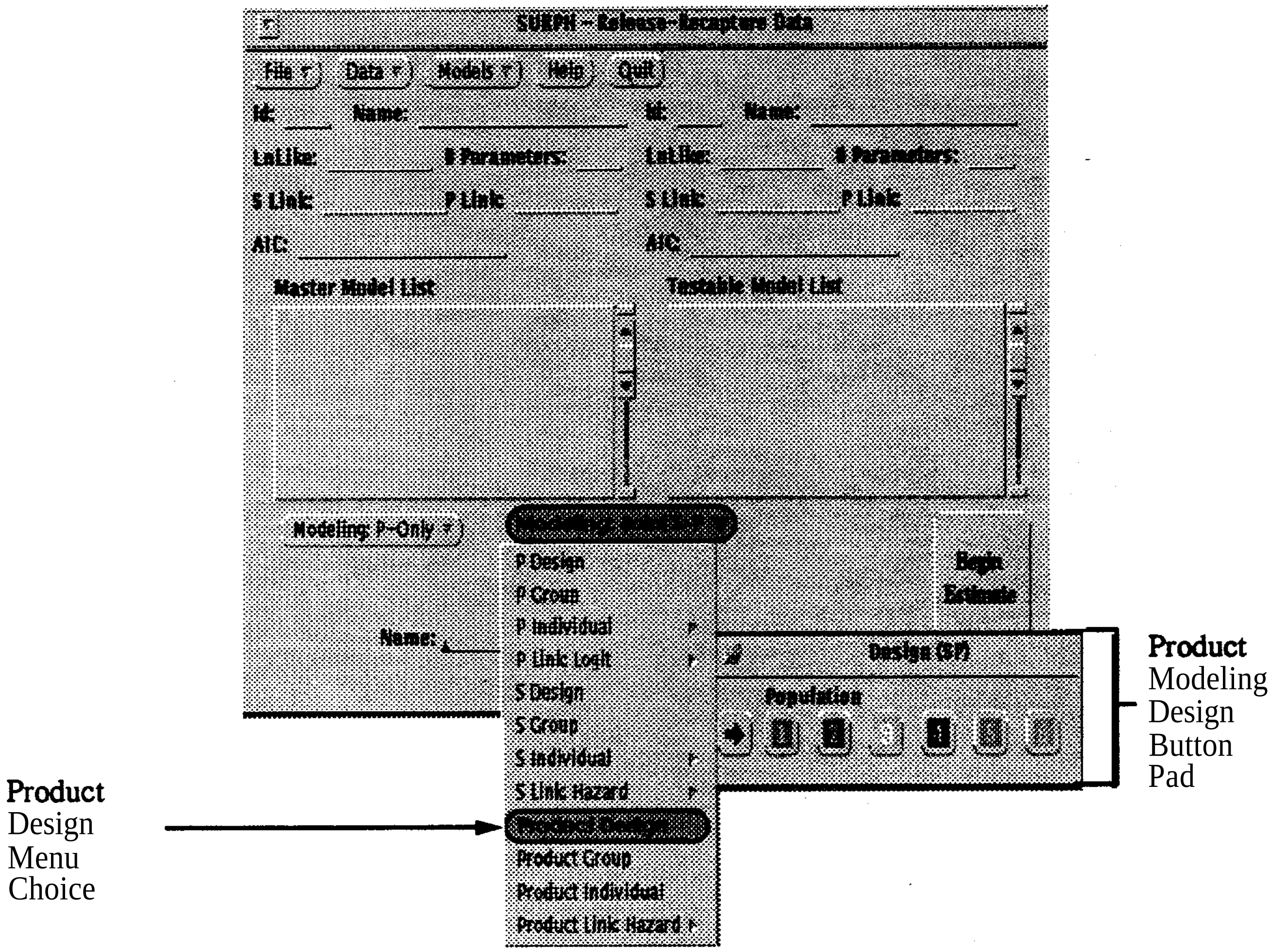
The example above shows a survival model with a common length effect for all populations in the 1st period (all “l’s” in the first column), a different common length effect for all populations in the 2nd period (all **“2’s”** in the second column), and no length effect elsewhere (all “O’s” in the third **column).**



S Link
Menu
Choice

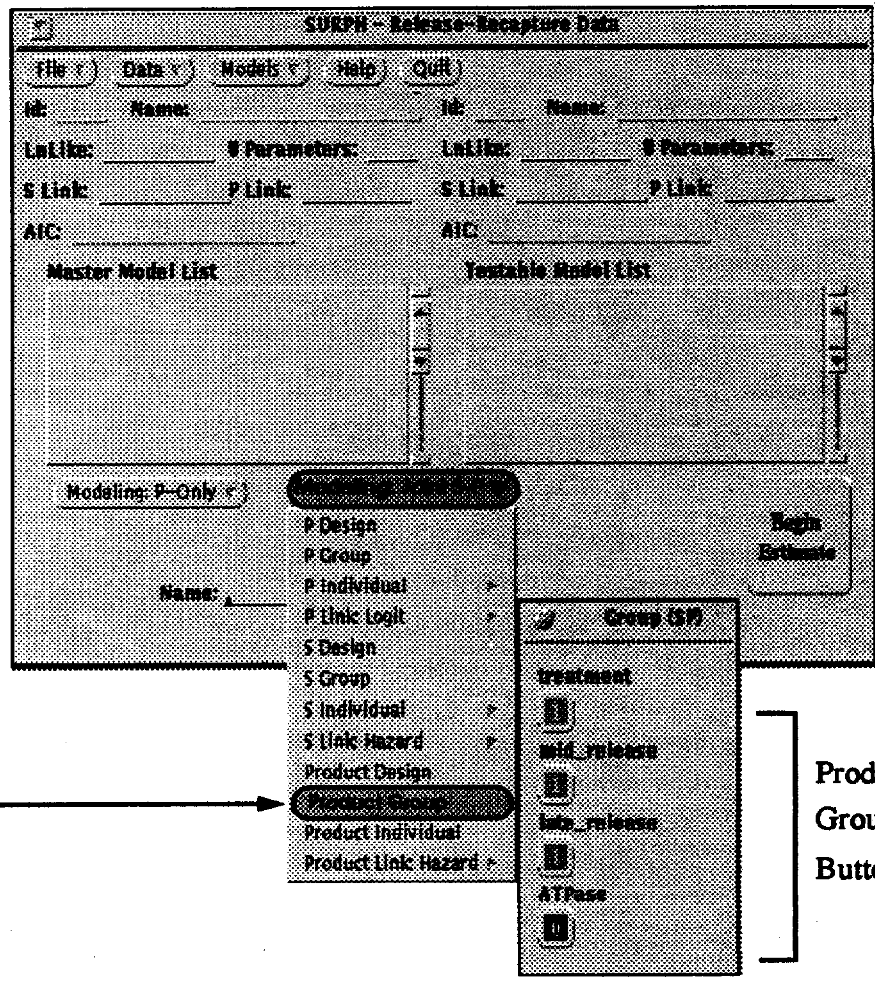
The Survival Modeling Link Function **(Modeling: Joint S-P -> S Link) allows** you to specify whether effects on the survival process will be linked to survival probabilities using a Hazard Function link or a **Logit** Function link. The default link for survival modeling is Hazard.

THIS PAGE INTENTIONALLY LEFT BLANK



In the final recapture occasion, although it is not possible to separately estimate capture and survival effects, it is possible to estimate effects on the product (Survival x Capture). The Product Design Button Pad **(Modeling: Joint S-P -> Product Design)** allows you to model the final period product as a function of population-specific effects. You change the **parameterization** of the product model by changing the button settings **within the Product Modeling Design Button Pad.** Unique button settings within the population row indicate unique effects for each population. Duplicated button settings within the population row indicate shared parameters for **multiple** populations. For the example above, all population effects are estimated with unique parameters, as each button **in the Product Modeling Design Button Pad is** set uniquely.

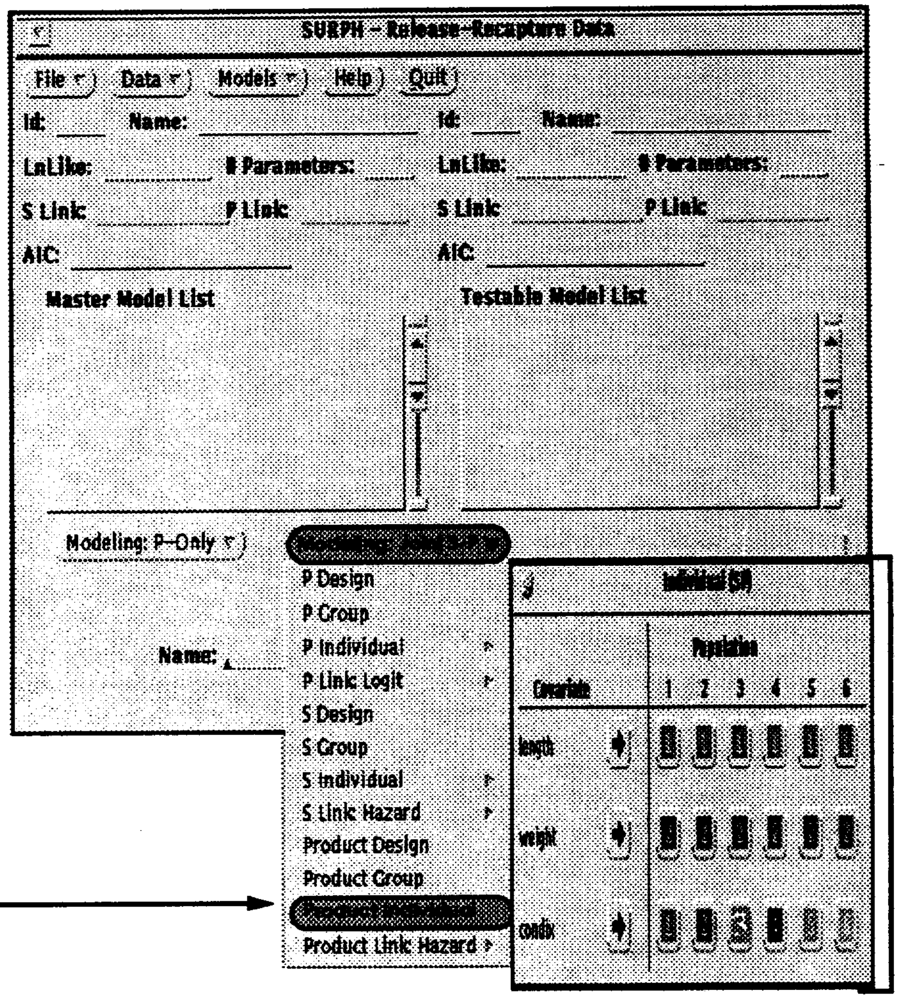
THIS PAGE INTENTIONALLY LEFT BLANK



The Product Group Button Pad **(Modeling: Joint S-P -> Product Group)** allows you to model the final period product (Survival x Capture) probabilities as a function of Group Covariates. You change the **parameterization** of the product model by altering the **buttton** settings within the **Product Modeling Group Covariate Button Pad.** A zero indicates that no **effect** is estimated for that wvariate, whereas a one indicates that a parameter is estimated for the effect of that covariate. Left-click on the individual buttons or use the Quick Buttons (see Quick Buttons - Instructions **) to change button settings.**

The model specified above shows a product model with effects of treatment, mid-release, and late-release wvariates (a "1" for each wvariate). No parameter is estimated for the effect of **ATPase** (a "0" beneath the **ATPase** covariate). Note that the treatment, mid-release and **late_release** settings are independent (i.e. the "1" settings do not indicate that the same parameter is used for all three **covariates).**

THIS PAGE INTENTIONALLY LEFT BLANK



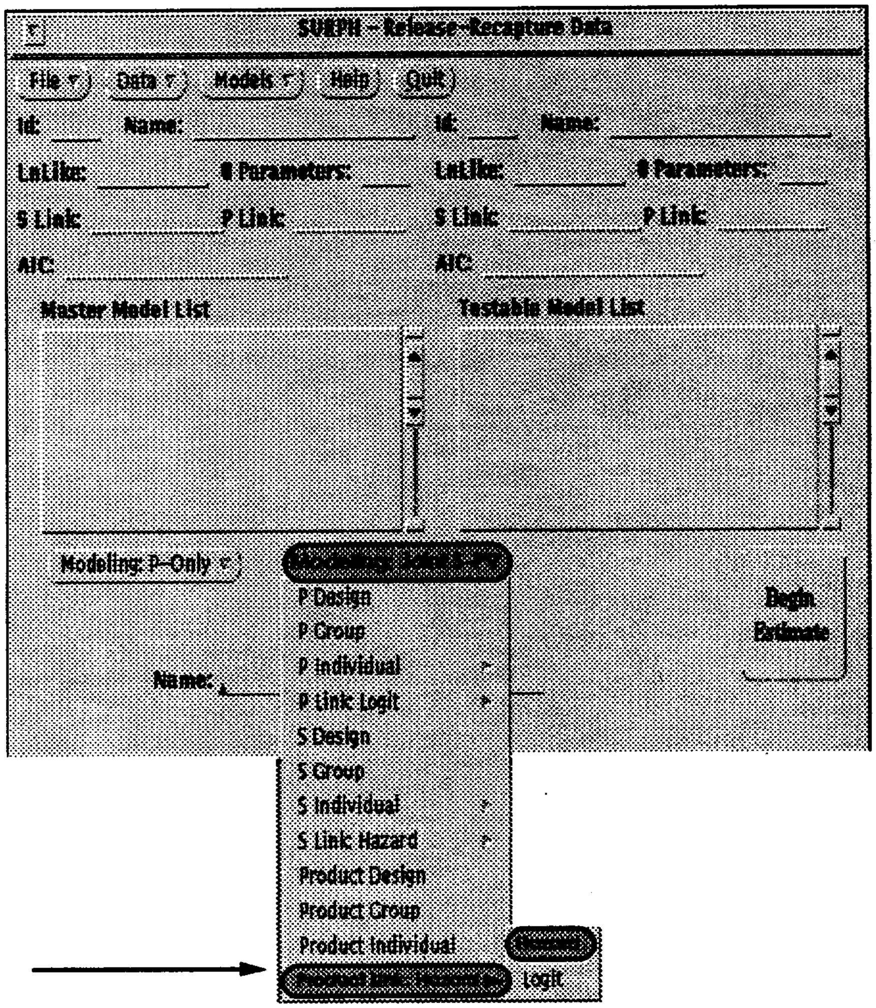
Product Individual Menu Choice

Product Modeling Individual Button Pad

The Product Modeling Individual Covariate Button Pad **(Modeling: Joint S-P -> Product Individual)** allows you to model the final period product (Survival x Capture) probabilities as a function of Individual Covariates. You access the **Product Modeling Individual Button Pad** by left-clicking on the Product Individual Menu Choice. You change the **parameterization** of the product model by changing the button settings **within the Product Modeling Individual Button Pad.** A zero indicates that no effect is estimated for that wvariate, whereas non-zero numbers indicate that a parameter is estimated for the effect of that wvariate. If the same button setting is used for a particular covariate, a common parameter is estimated for the effect of the wvariate for all populations sharing that setting. If different button settings are used, separate parameters are estimated. Left-click on the individual buttons or use the Quick Buttons to change button settings (see Quick Buttons - Instructions **).**

The model specified above shows a product model with a **common** length effect for all populations (all "1 **'s"),** a common weight effect for all populations (all "2's"), and unique wndix (wndition index) effects for each population (Button settings are "123456").

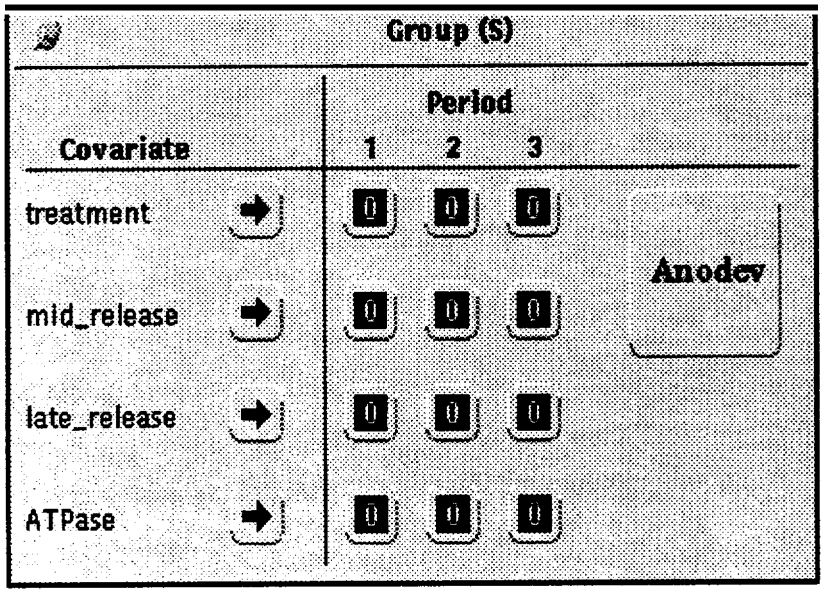
Note that the length, weight and condix settings are independent. Thus, if a common length effect for all populations had been parameter&d with all "l's", and a common weight effect for all populations had been parameter&d with all "1 's", the model would estimate a common effect for weight, and a separate wmmon effect for length. The model would be no different from the model shown above, where the wmmon length effect for all populations is indicated by all "1 's", and the wmmon weight effect for all populations is indicated by all "2's".



Product Link
Menu Choice

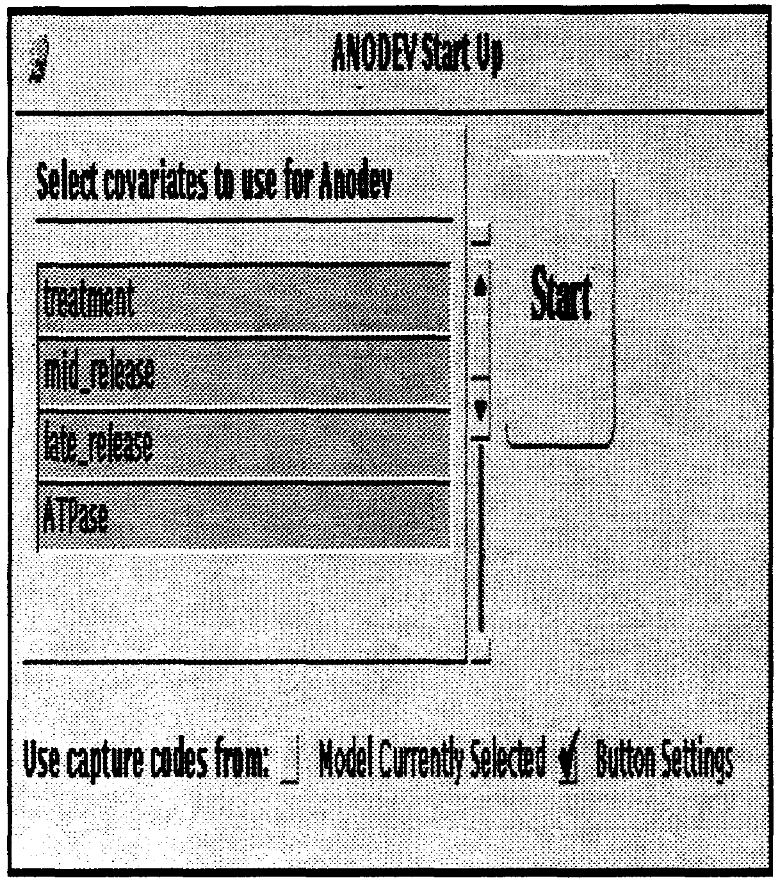
The Product Modeling Link Function **(Modeling: Joint S-P -> Product Link)** allows you to specify whether effects on the final period product (Survival x Capture) will be linked to the product probabilities using a Hazard Function link or a Logit Function link. The default link for final product modeling is Hazard.

THIS PAGE INTENTIONALLY LEFT BLANK



This Button Pad **(Modeling: Joint S-P -> S Group) allows the** user to model group covariate effects on survival. The user can test the **significance** of the group covariate effects using the likelihood ratio approach (see Section 2.4.1 and Survival Modeling: Group Covariates **) or using the Analysis of Deviance approach (see** Section 2.4.3). To analyze the data using Analysis of Deviance, left-click on the **Anodev Button.** SURPH **will** prompt the user to select the group covariates to analyze (see ANODEV Start Up **),** and will then run three default models required for ANODEV (the Full Model, the Period Means Model, and the Total Model). An Analysis of Deviance Table will appear when all calculations are completed (see all pages associated with Analysis of Deviance).

THIS PAGE INTENTIONALLY LEFT BLANK

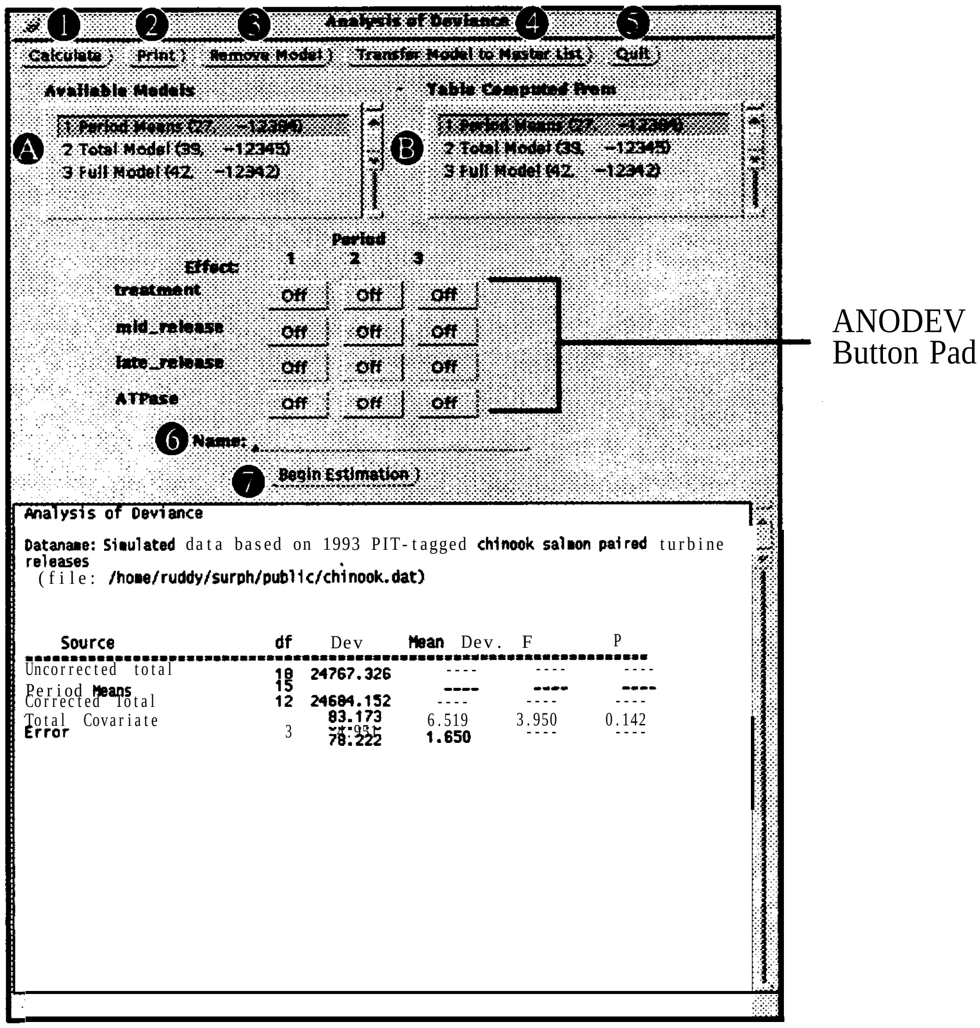


This pop-up window appears after you initiate the Analysis of Deviance option **(Modeling: Joint S-P -> S Group -> Anodev). You** must select the group covariates that will be used during the Analysis of Deviance. You select a group wvariate by left-clicking on the group covariate name in the list beneath "Select covariates to use for ANODEV". Group covariates that have been selected will appear darkened. You remove a group **covariate** from those that will be used during the Analysis of Deviance by left-clicking on a selected wvariate name.

You must also select the capture model that is to be used during the analysis. If the "Model Currently Selected" **checkbox** is checked, the model for capture probabilities during ANODEV is the same as that in the model currently selected in the **Master** Model List in the Base Window. If the "Button Settings" **checkbox** is checked, the capture model is read from the current settings on the Button Pads for capture probabilities.

After selecting the group **covariates** to analyze, and the model for capture probabilities, left-click on the **Start Button** 'to initiate Analysis of Deviance. When the start button is pushed, the three "standard models" (Full, Period Means, and Total Covariate Models) are estimated (see Sections 3.5 and 4.5 in the SURPH.l manual). **After** the standard models are complete, the Analysis of Deviance Table appears (see

Analysis of Deviance Table (A)



This window **(Modeling: Joint S-P -> S Group -> ANODEV)** contains a control panel (top) for guiding Analysis of Deviance and a text pane (bottom) for displaying ANODEV results. The main features of the ANODEV Table are the following:

Ⓐ **Available Models List -** A list of models available for potential use in computing ANODEV Tables. All models in this list have the model for capture probabilities and involve the group wvariate effects on survival selected at the ANODEV start-up.

Ⓑ **Table Computed From List -** A list of models that you selected from the **Available Models List** to be used to compute an ANODEV Table. This list always contains the Full, Period Means, and Total Covariate Models. The other models in the list form a hierarchical set that is used to partition the Total Covariate Deviance. A model is added to **the Table Computed From List** by selecting (left-clicking) on the model in the

Available Models List. If the model selected fits in the hierarchy currently defined by the models **in the Table Computed From List, the** model **is added to the Table Computed From List.** If a new hierarchy is desired, models must **first** be removed from the **Table Computed From List (see Remove Model Button).**

❶ **Calculate Button -** Left-click to compute the ANODEV Table from the models currently **in the Table Computed From List.**

❷ **Print Button -** Left-click to send a copy of the current ANODEV Table to the printer.

❸ **Remove Model Button -** Left-click to remove the currently selected model from **the Table Computed From List.** Models must **be** removed from **the Table Computed From List if an** alternative hierarchy (partition of the Total Deviance) is to be investigated.

❹ **Transfer Model Button -** Left-click to transfer the currently selected model from **the Available Models List** to **the Master Model List** on **the SURPH Base Window.** While they are legitimate SURPH models, the three standard models and any models estimated from the ANODBV Button Pad (see below) are by default not listed on the **Master Model List** on the SURPH Base Window. On the basis of ANODEV results, it is frequently desirable to transfer selected models from **the** ANODEV **Available Models List** to the **Master Model List,** for functions not available from the ANODEV Table (e.g., display of parameter estimates; residual plots; individual covariate modeling).

❺ **Quit Button -** Left-click to quit Analysis of Deviance. This is the only way to leave ANODEV; the pushpin in the left-hand corner of the ANODEV Table is disabled. You are asked to confirm the intent to quit ANODEV **(see** Analysis of Deviance Table (C)).

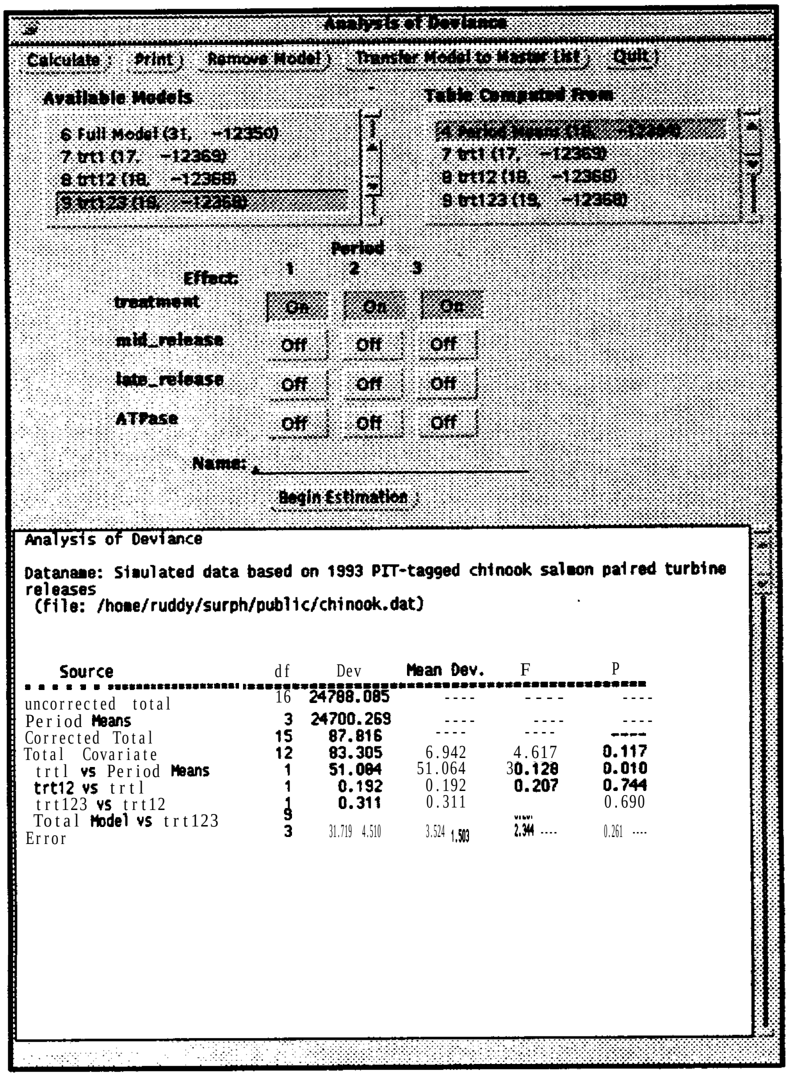
❻ **Model Name Field -** Left-click on the horizontal line to place the cursor in the field, then type the name for a new model.

❼ **Begin Estimation Button -** Left-click to fit the model **currently** defined by the **ANODEV Button Pad. If the model has already been estimated, a warning appears (see** Analysis of Deviance Table (B) **). After estimation is complete, the model is added to** the Available Models List. If a name for the model is not specified in the **Model Name Field, the** model's name is "Model 'n"', where **'n'** is the model number.

ANODEV Button Pad - While performing Analysis of Deviance, new models may be estimated to partition the Total **Covariate** Deviance into contributions of **specific** covariates **in** each **period.** ANODEV models **are specified** on **the ANODEV Button Pad** rather **than the usual Survival- Group Covariate Button Pad. The ANODEV Button Pad** functions **similar to the Group Covariate Button Pad,** except that it is not possible to define equalities among parameters across periods.

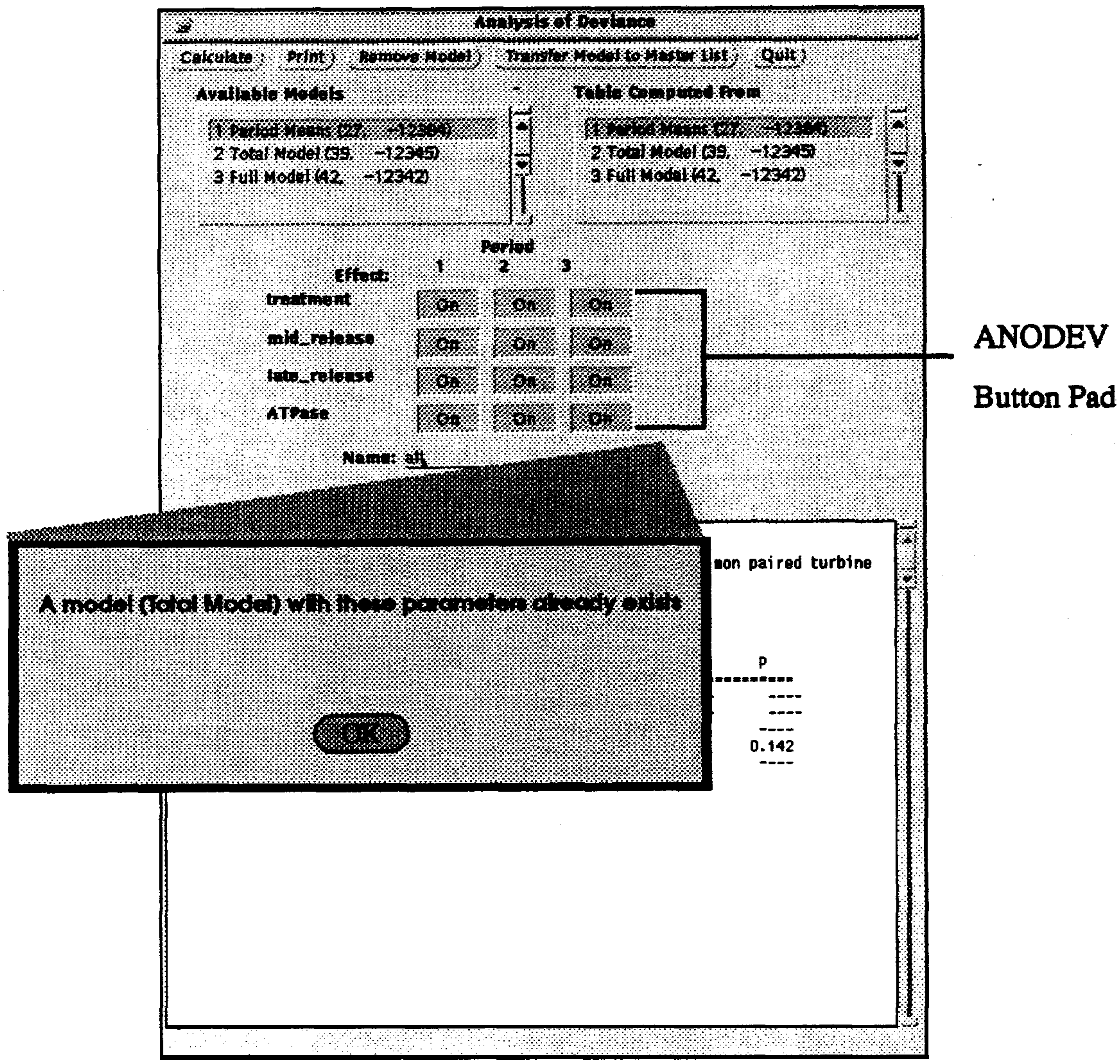
ANODEV Text Pane - Displays the results of Analysis of Deviance. The **Dataname** and file from which the data were read are listed at the top of the pane, followed by the Analysis of Deviance table, based on the models in the current **Table Computed From List (see** Section 4.5 in the SURPH.l manual for a full explanation of the ANODEV Table). When the ANODEV Table **first** appears, the **displayed** ANODEV Table is computed from the Full, Period Means, and Total Covariate **Models** (i.e., the Total Covariate Deviance is displayed, but it is not partitioned in any way).

THIS PAGE INTENTIONALLY LEFT BLANK



Analysis of Deviance can be used to test the significance of group covariate effects on survival. New models that partition the Total Covariate Deviance into contributions of specific covariates in each period can be created by adding (On) or deleting (Off) parameters using the ANODEV Button Pad. After a series of hierarchical models have been estimated, you can create an ANODEV Table. Left-click on model names in the **Available Models List** to move them to the **Table Computed From List.** Then, to create the ANODEV Table, left-click on the **Calculate Button.**

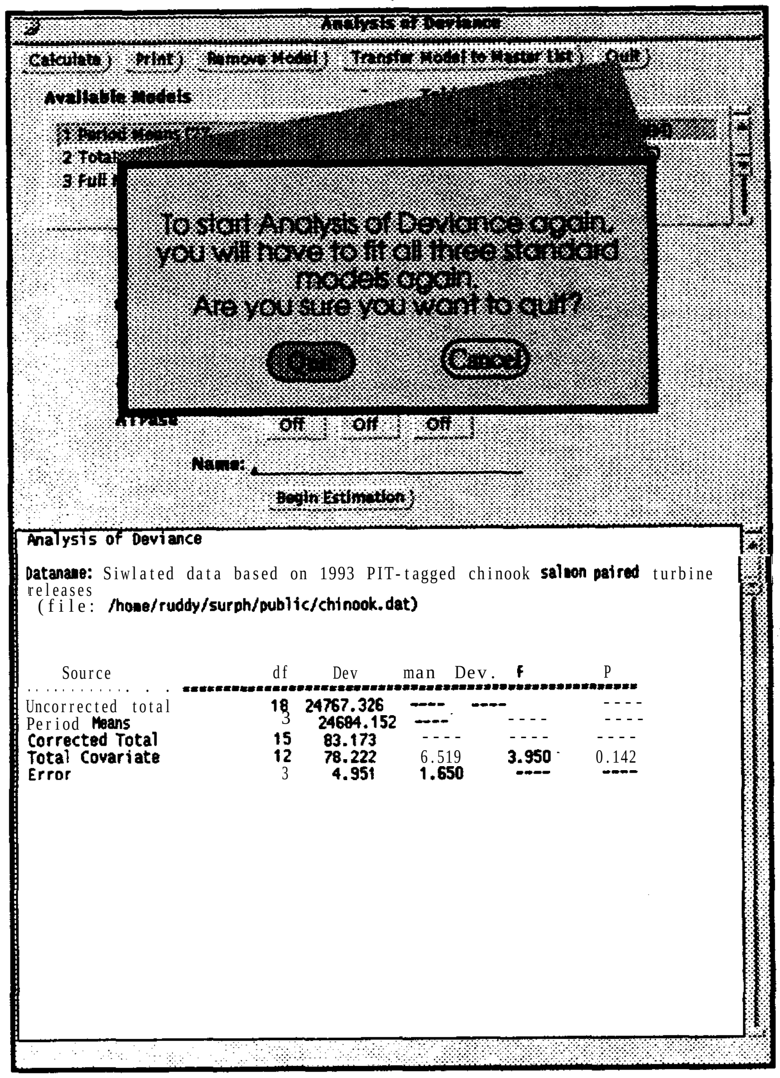
In the example above we have estimated three nested models (labelled above as trtl, trt12, trt123). Model "trt1" has a treatment effect estimated in period 1, model "trt12" has separate treatment effects estimated in periods 1 and 2, and model "trt123" has separate treatment effects estimated in periods 1, 2, and 3. You read the ANODEV Table as you would an ANOVA Table. Thus, the only significant group covariate effect from this set of nested models is a treatment effect in period 1 (p=0.01). Recall that due to non-orthogonality, the treatment effect in subsequent periods may be affected by whether the treatment effect in the first period has been included in the model. Therefore, you may also wish to estimate the treatment effects in periods 2 and 3 independent of the effect in period 1.



When an ANODEV model that has already been estimated is specified, a message window informs you of this duplication.

In the example, using the Modeling Buttons, we have attempted to run the Total Covariate Model under an alternative name. The message window informs us that this model already exists.

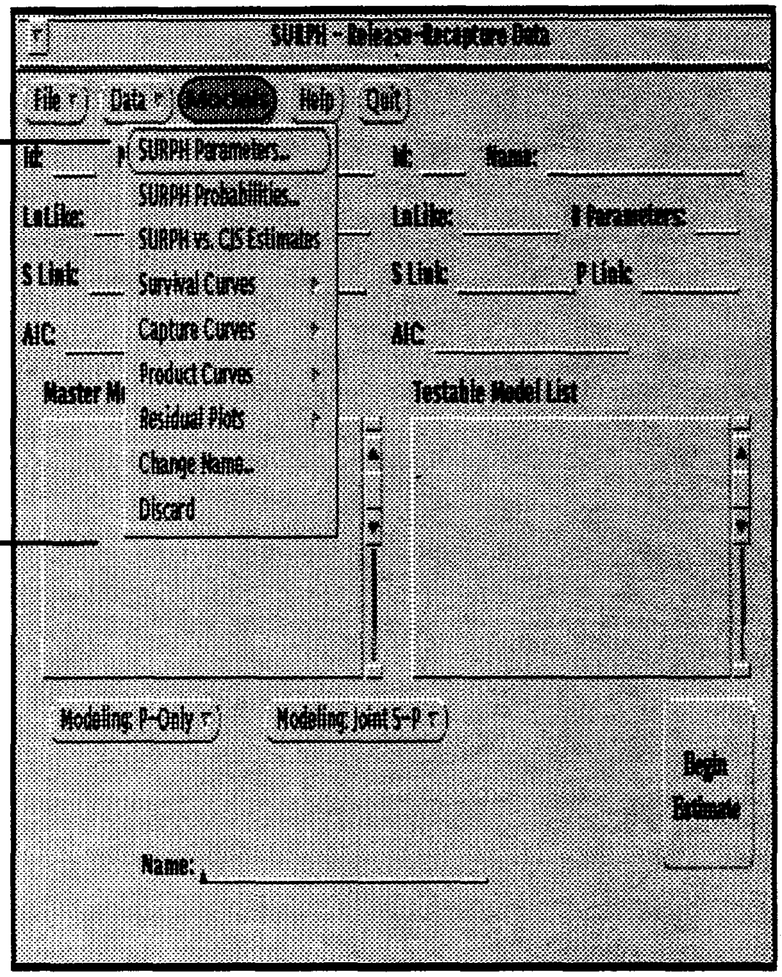
THIS PAGE INTENTIONALLY LEFT BLANK



When you attempt to quit the **ANODEV** window, a pop-up message informs you that the three standard models will be lost if ANODEV is quit. As the estimation of the three standard models can take several minutes, you may wish to reconsider.

THIS PAGE INTENTIONALLY LEFT BLANK

Models
Pull-Down
Menu



This is the pull-down menu **accessed** by right-clicking **on the Models Button. The** Model Pull-Down Menu has **choices** to display **SURPH** survival and capture estimates, view graphics that illustrate group and individual covariate effects on survival and capture probabilities, view residual plots to investigate goodness-of-fit, change model names, and discard models. All actions pertain **to** the model currently selected in the **Master Model List.**

SURPH Parameters - Display text window with estimated parameter values, standard errors of parameter **values,** and the variance-covariance matrix for the model currently selected in the **Master Model List** **(see** SURPH Parameters **).**

Estimates are displayed on the scale on which they are **estimated** by SURPH (e.g., β in the expression $S^{exp(X'\beta)}$), rather than as **probabilities**.

SURPH Estimates - Displays text window with "integrated probability estimates" for the model currently selected in the **Master Model List. Estimates are displayed** on the probability scale (e.g., $\hat{S}_{ik} = S^{exp(X'_{ik}\beta)}$ (see SURPH Probabilities).

SURPH vs. CJS Estimates - Displays plot of integrated probability estimates of capture and survival probabilities versus the **corresponding Cormack/Jolly-Seber** estimates (see SURPH vs. CJS Estimates).

Survival Curves - Displays graphical representations of group and individual covariate effects on survival probabilities or on relative risk (see Survival Graphics and Relative Risk Graphics).

Capture Curves - Displays graphical representations of group and individual covariate effects on capture probabilities (see Capture Graphics).

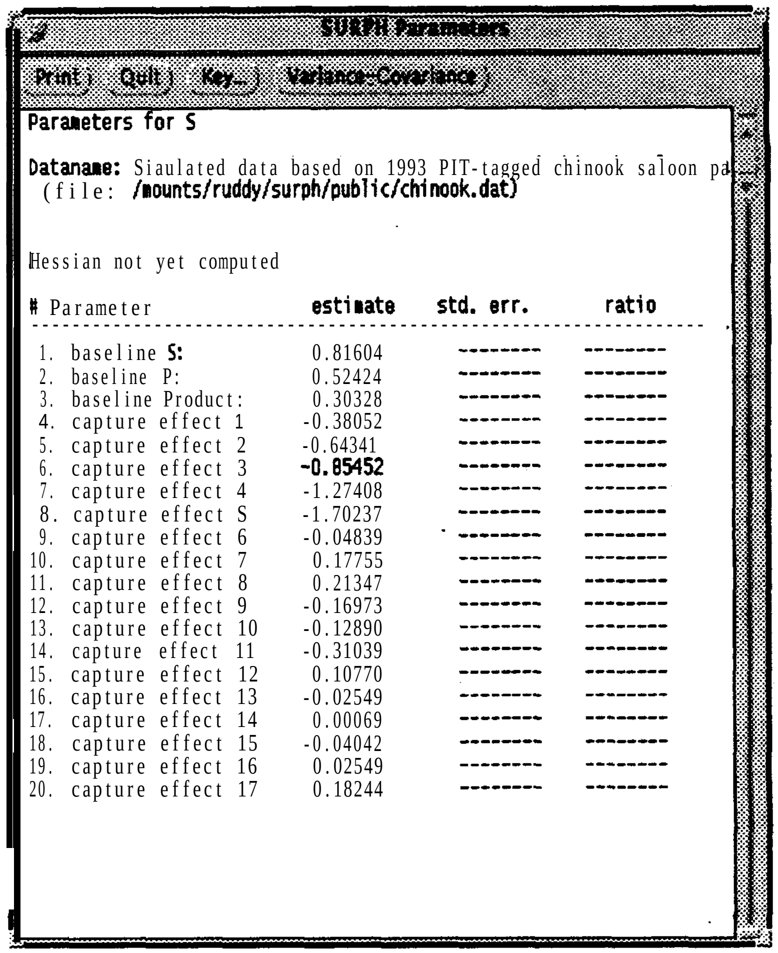
Product Curves - Displays graphical representations of group and individual covariate effects on the final period product (Survival x Capture) probabilities (see Product Graphics).

Residual Plots - Displays graphical representations of model residuals from two perspectives, a) Residuals vs. Capture History, or b) Normal Q-Q Plot (see Residual Plot or Quantile-Quantile Plot).

Change Name - Changes the name of **the** currently selected model **in the Master Models List (see** Change Model Name

Discard - Discards the currently selected model in the **Master Model** Lii. Models that are discarded are no longer accessible to you and have **to be recomputed if** needed **again(see** Confirm Discard Dialog Box **).**

THIS PAGE INTENTIONALLY LEFT BLANK



SURPH Parameters

Print | Quit | Key... | Variance-Covariance

Parameters for S

Dataname: Siaulated data based on 1993 PIT-tagged chinook saloon pa
(file: **/mounts/ruddy/surph/public/chinook.dat)**

Hessian not yet computed

# Parameter	estimate	std. err.	ratio
1. baseline S:	0.81604	--------	--------
2. baseline P:	0.52424	--------	--------
3. baseline Product:	0.30328	--------	--------
4. capture effect 1	-0.38052	--------	--------
5. capture effect 2	-0.64341	--------	--------
6. capture effect 3	-0.85452	--------	--------
7. capture effect 4	-1.27408	--------	--------
8. capture effect S	-1.70237	--------	--------
9. capture effect 6	-0.04839	--------	--------
10. capture effect 7	0.17755	--------	--------
11. capture effect 8	0.21347	--------	--------
12. capture effect 9	-0.16973	--------	--------
13. capture effect 10	-0.12890	--------	--------
14. capture effect 11	-0.31039	--------	--------
15. capture effect 12	0.10770	--------	--------
16. capture effect 13	-0.02549	--------	--------
17. capture effect 14	0.00069	--------	--------
18. capture effect 15	-0.04042	--------	--------
19. capture effect 16	0.02549	--------	--------
20. capture effect 17	0.18244	--------	--------

This text window displays values of the estimated parameters for the model currently selected **in the Master Models List (Models -> SURPH Parameters).** Estimates are displayed on the scale on which they are estimated by **SURPH** (e.g., "capture effect 1" may be the second period effect (ρ_2) such that the capture probability **in period 2 is** $P^{exp(\rho_2)}$.

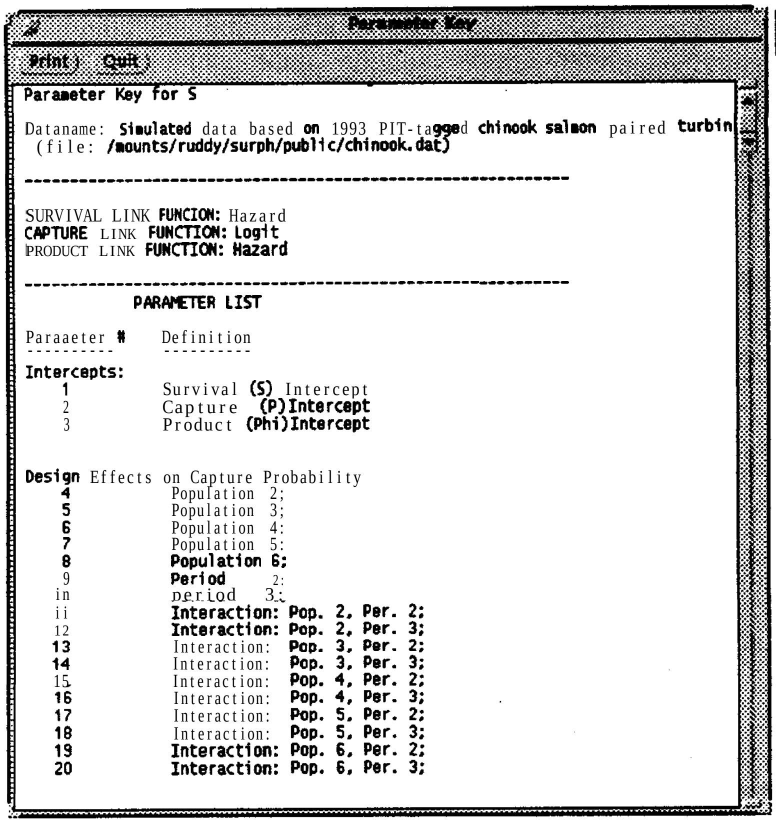
There are also columns for estimated standard errors and the ratios of parameters to standard errors. The first time the SURPH parameters text window is displayed for a particular model, these columns are blank because the variance-covariance matrix is not calculated until you left-click on **the Variance-Covariance Button. The** numerical procedure to estimate the variance-covariance matrix can take several minutes or more on some computers, and you **will** sometimes find it handy to view the parameter estimates alone, without waiting for the variance-covariance computation.

Print Button - Left-click to send a copy of this text window information to the printer.

Quit Button - Left-click to dismiss this text window.

Key Button - Left-click to view the key to the parameter list and two distinct cross-references.

Variance-Covariauce Button - Left-click to calculate and display the estimated Variance-Covariance Matrix of the parameter estimates.



```
Parameter Key for S

Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbin
 (file: /mounts/ruddy/surph/public/chinook.dat)
-----------------------------------------------------------

SURVIVAL LINK FUNCION: Hazard
CAPTURE LINK FUNCTION: Logit
PRODUCT LINK FUNCTION: Hazard
-----------------------------------------------------------
           PARAMETER LIST

Paraaeter #     Definition
---------       ----------
Intercepts:
   1            Survival (S) Intercept
   2            Capture  (P)Intercept
   3            Product (Phi)Intercept

Design Effects on Capture Probability
   4              Population  2;
   5              Population  3;
   6              Population  4:
   7              Population  5:
   8              Population 6;
   9              Period      2:
   in             period      3:
   ii             Interaction: Pop. 2, Per. 2;
   12             Interaction: Pop. 2, Per. 3;
   13             Interaction: Pop. 3, Per. 2;
   14             Interaction: Pop. 3, Per. 3;
   15             Interaction: Pop. 4, Per. 2;
   16             Interaction: Pop. 4, Per. 3;
   17             Interaction: Pop. 5, Per. 2;
   18             Interaction: Pop. 5, Per. 3;
   19             Interaction: Pop. 6, Per. 2;
   20             Interaction: Pop. 6, Per. 3;
```

A

This text window **(Models -> SURPH Parameters -> Key)** provides a key to the survival and capture parameter list for the model that is currently selected in the Master **Model List.** Buttons across the top row **allow you** to **Print the** information contained within **this** window or **allow you** to **Quit (dismiss) this** window.

There are three portions to the key. You move from one portion to the next by scrolling through the text pane. The first portion of the key (top, ❑) displays the Model Name **("S"** in this case), **Dataname** (chinook&) and the Survival, Capture and Product Link Functions. This is followed by a listing of the estimated parameters, their numeric designations and their definitions. In the example text window (top), parameter 2 is the Capture Intercept Parameter, whereas parameter 9 is the **2nd** period effect on capture.

The second portion of the key (middle, ❑) displays the Parameter Cross-Reference 1. This listing indicates, by population and period, which parameters (by numeric designation) are involved in the estimation of capture, survival, and product probabilities. In the example text pane (middle), notice that survival for all populations and periods is affected solely by the **1st** parameter (Survival Intercept). This is because the **Model S (see** Survival Modeling - Examples **) was used to estimate survival probabilities. However, since the Model Pst was used to estimate capture probabilities,** each population and period has a unique suite of capture parameters. For example, for the **2nd** population in the **2nd** period, the capture parameters used to estimate the capture probability are parameters 2 (Capture **Intercept)**, 4 (effect of the **2nd** population on capture), 9 (effect of the **2nd** period on capture), and 11 (interaction of the **2nd** population and **2nd** period effects on capture).

Parameter Key

Print Quit

B

PARAMETER CROSS-REFERENCE 1

Parameters Affecting Survival Probability:

Population	Period	Parameter
1	1	1
1	2	1
1	3	1
2	1	1
2	2	1
2	3	1
3	1	1
3	2	1
3	3	1
4	1	1
4	2	1
4	3	1
5	1	1
5	2	1
5	3	1
6	1	1
6	2	1
6	3	1

Parameters Affecting Capture Probability:

Population	Period	Parameter
1	1	2
1	2	2, 9
1	3	2, 10
2	1	2, 4
2	2	2, 4, 9, 11
2	3	2, 4, 10, 12
3	1	2, 5
3	2	2, 5, 9, 13
3	3	2, 5, 10, 14
4	1	2, 6
4	2	2, 6, 9, 15
4	3	2, 6, 10, 16
5	1	2, 7
5	2	2, 7, 9, 17
5	3	2, 7, 10, 18
6	1	2, 8
6	2	2, 8, 9, 19
6	3	2, 8, 10, 20

The final portion of the key (bottom, C) displays the Parameter Cross-Reference 2. This display, rather than being organized by populations and periods, is **organized** by Design, Group, and Individual Effects on Survival, Capture and Product Probabilities. This portion of the key is used to understand the Design, Group Govariate, and Individual Covariate Effects on Populations and Periods. For example, for the **2nd** population in the **2nd** period, the capture parameters used to estimate the capture probability are the Capture Intercept, which is included in all capture probabilities (parameter **2),** the Design **Effect** on Capture of the **2nd** population (parameter **4),** the Design Effect on Capture of the **2nd** period (parameter **9),** and the Design Effect on Capture of the **2nd** population and **2nd** period interaction (parameter 11). All Group and Individual Covariate **Effects** on capture are **zero,** which indicates that the fitted model ("Pst") did not include these effects.

C

Parameter Key

Print Quit

```
------------------------------------------------------------
Design Effects on Capture Probability

Population         1   1   2   3   4   5   6
-------------------+--------------------------
Paraaeter          |   0   4   5   6   7   8

Period             I   ?   2   3
-------------------+-------------
Parameter          I   D   9   10

                       Period
Interactions       |   1   2   3
-------------------+-------------
                 1 |   0   0   0
Population       2 |   0  11  12
                 3 |   0  13  14
                 4 |   0  15  16
                 5 |
                 6 | 0 0  17 19  18 20

Croup Covariate Effects on Capture Probability

Period             |   1   2   3
-------------------+-------------
t reatment             0   0   0
● ihreleas         I   0   0   0
late_relea         |   0   0   0
ATPase             |   0   0   0

Individual Covariate Effects on Capture Probability

                       Period
length             |   1   2   3
-----------------1 |-------------
Populatlon       2 |   0   0   0
                       0   0   0
                 3 |   0   0   0
                 4 |   0   D   0
                 5 |   0   0   0
                 6 |   0   0   0
```

THIS PAGE INTENTIONALLY LEFT BLANK

SURPH Parameters

Print Quit Key.. Variance-Covariance

Parameters for S

Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
(file: /mounts/ruddy/surph/public/chinook.dat)

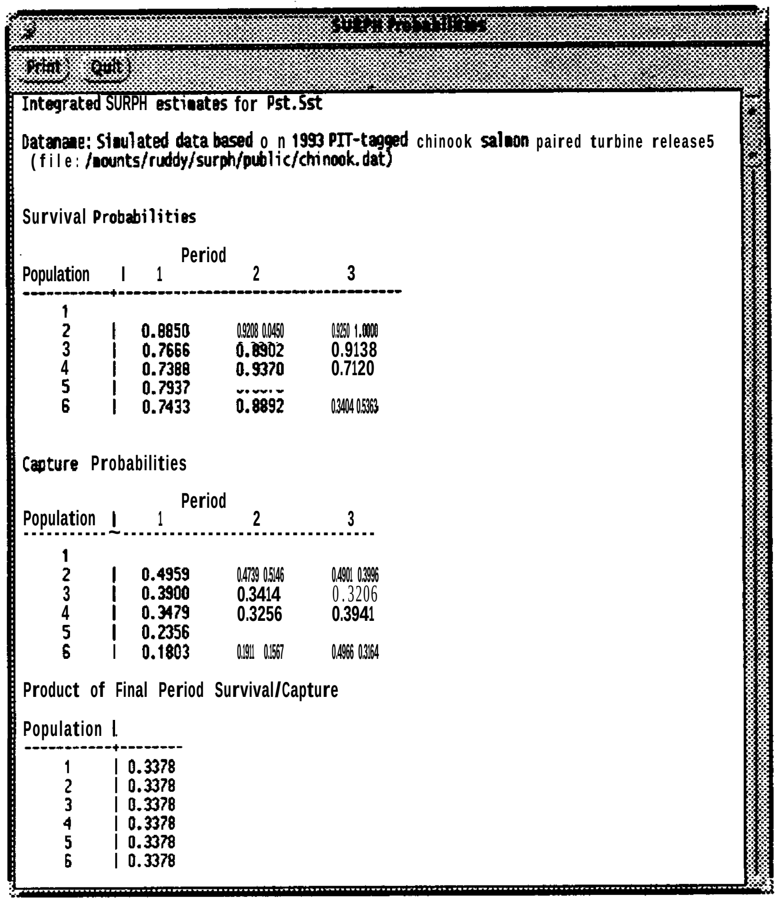
# Parameter	estimate	std. err.	ratio
1. baseline S:	0.81604	0.00719	113.52921
2. baseline P:	0.52424	0.01602	32.72289
3. baseline Product:	0.30328	0.01177	25.76787
4. capture effect 1	-0.38052	0.09982	-3.81194
5. capture effect 2	-0.64341	0.09942	-6.47143
6. capture effect 3	-0.85452	0.10140	-8.42753
7. capture effect 4	-1.27488	0.08417	-15.14702
8. capture effect 5	-1.70237	0.11562	-14.72430
9. capture effect 6	-0.04839	0.07662	-0.61553
10. capture effect 7	0.17755	0.11026	1.61017
11. capture effect 8	0.21347	0.13143	1.62420
12. capture effect 9	-0.16973	0.15959	-1.06350
13. capture effect 10	-0.12890	0.13223	-0.97482
14. capture effect 11	-0.31039	0.15915	-1.95834
15. capture effect 12	0.10770	0.13419	0.80258
16. capture effect 13	-0.02549	0.16099	-0.15836
17. capture effect 14	0.00069	0.03841	0.01797
18. capture effect 15	-0.04042	0.15526	-0.26034
19. capture effect 16	0.02549	0.15531	0.16412
20. capture effect 17	0.18244	0.17564	1.03873

Variance-covariance matrix:

	1	2	3	4	5	6	7	8
1	0.0001	-0.0000	-0.0001	-0.0001	-0.0000	-0.0000	-0.0001	-0.0000
2	-0.0000	0.0003	0.0000	-0.0010	-0.0010	-0.0010	-0.0007	-0.0010
3	-0.0001	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
4	-0.0000	-0.0003	0.0000	0.0100	0.0042	0.0041	0.0029	0.0041
5	-0.0000	-0.0003	0.0000	0.0042	0.0099	0.0041	0.0029	0.0041
6	-0.0000	-0.0003	0.0000	0.0041	0.0041	0.0103	0.0029	0.0041
7	-0.0000	-0.0002	0.0000	0.0029	0.0029	0.0029	0.0071	0.0029
8	-0.0000	-0.0003	0.0000	0.0041	0.0041	0.0041	0.0029	0.0134

This illustration **(Models -> SURPH Parameters -> Variance-Covariance)** shows the SURPH Parameters text window as it appears after left-clicking on the **Variance-Covariance Button.** It is sometimes necessary to expand the window to view the matrix elements. This estimated variance-covariance is based upon the Cramer-Rao Lower Bound (Hogg and Craig 1978) under Maximum Likelihood theory. Each element on the diagonal of the matrix is the estimated variance for the associated parameter (e.g. the estimated variance for parameter **#3** - baseline Product - is the element **(3,3)** of the Variance-Covariance Matrix, whose value is 0.0001). Off-diagonal elements are the estimated covariance between the two associated parameters (e.g., the estimated covariance between parameters **#3** and **#1** - baseline Product and baseline S - is the element **(3,1)** of the Variance-Covariance Matrix, whose value is -0.0001). All diagonal elements should be non-negative, as variances are strictly non-negative. Off-diagonal elements may be negative or positive.

Once the variance-covariance matrix has been estimated, the **standard** error and the ratio of the estimate to its standard error is &played in the upper region of the text pane. **The** ratio of $\hat{\theta}/\left(\sqrt{var(\hat{\theta})}\right)$ is asymptotically distributed as a standard normal deviate (i.e., N(O,l)).



SURPH Probabilities

Print Quit

Integrated SURPH estimates for Pst.Sst

Dataname: Simulated data based o n 1993 PIT-tagged chinook salmon paired turbine release5
(file: /mounts/ruddy/surph/public/chinook.dat)

Survival Probabilities

Population	Period 1	2	3
1			
2	0.8850	0.9208 0.0450	0.9250 1.0000
3	0.7666	0.8902	0.9138
4	0.7388	0.9370	0.7120
5	0.7937	[illegible]	
6	0.7433	0.8892	0.3404 0.5363

Capture Probabilities

Population	Period 1	2	3
1			
2	0.4959	0.4739 0.5146	0.4901 0.3996
3	0.3900	0.3414	0.3206
4	0.3479	0.3256	0.3941
5	0.2356		
6	0.1803	0.1911 0.1567	0.4966 0.3164

Product of Final Period Survival/Capture

Population	
1	0.3378
2	0.3378
3	0.3378
4	0.3378
5	0.3378
6	0.3378

This text pane **(Models -> SURPH Estimates)** provides survival and capture probability estimates for the model that is currently selected in the **Master Model List.** These probabilities are calculated from the SURPH parameters for the effects modeled. When the selected SURPH model does not include individual-based covariates, the probabilities are computed from the formula:

$$S_{ik} = S^{exp(\pi_i + \rho_k + (\pi\rho)_{ik} + \underset{\sim}{X}_{ik}\underset{\sim}{\beta}_k)}$$

where S_{ik}	is the survival probability for population ***i in period k***;
S	is the intercept parameter for survival;
π_i	is the effect for population ***i;***
ρ_k	is the effect for period ***k;***
$(\pi\rho)_{ik}$	is the effect for the population ***i/period k*** interaction;
$\underset{\sim}{X}_{ik}$	is the vector of group covariates for population ***i in period k*; and**
$\underset{\sim}{\beta}_k$	is the vector of slope parameters for group covariates in period k.

When the selected model includes individual-based covariates, the probabilities are "integrated" across all individuals **according** to the formula:

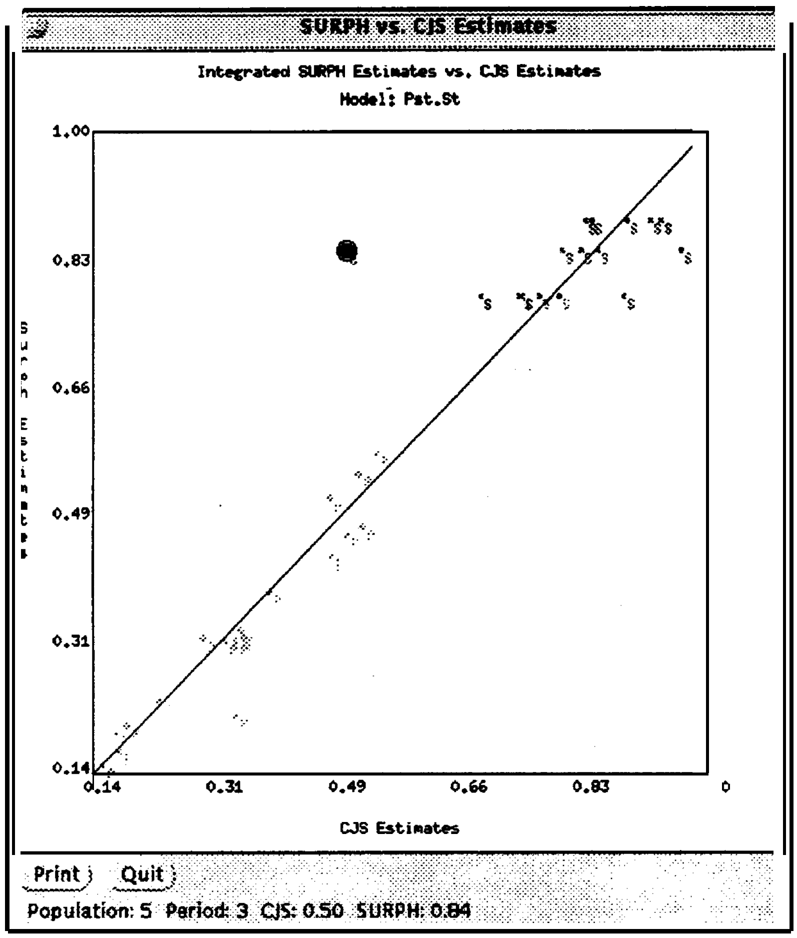
$$S_{ik} = \frac{\sum_{l=1}^{n_{ik}} S^{exp(\pi_i + \rho_k + (\pi\rho)_{ik} + \underset{\sim}{X}_{ik}\underset{\sim}{\beta}_k + \underset{\sim}{Y}_{il}\underset{\sim}{\gamma}_{ik})}}{n_{ik}}$$

where n_{ik} is the number of animals at risk in population ***i in period k***;

$\underset{\sim}{Y}_{il}$ is the vector of individual covariates for individual ***l*** in population ***i;***

and $\underset{\sim}{\gamma}_{ik}$ is the vector of slope parameters for the individual covariates for population ***i in period k.***

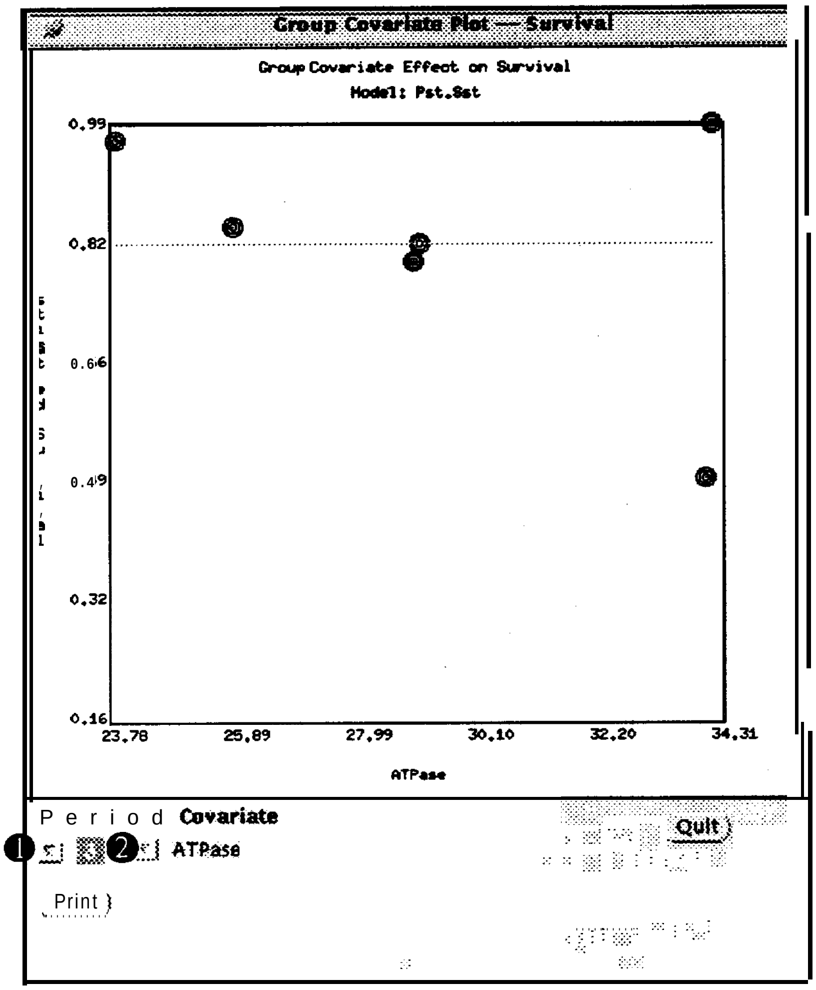
Buttons in the control panel at the top of the window allow you to Print the information contained within this window or allow you to Quit this window.



This window **(Models -> SURPH** vs. **CJS)** provides a graphical display of the "integrated" **SURPH** probability estimates for the model currently selected **in the Master Model List** versus the **Cormack/Jolly-Seber** (Cormack 1964, Jolly 1965, Seber 1965) estimates. Both Capture ("P's" on the graphic) and Survival ("S's" on the graphic) probabilities are plotted. The control panel along the bottom of the window has buttons to **Print** a copy of this window or to **Quit (dismiss) this** window.

This graphic provides a visual diagnostic for assessing the fit of models that have fewer parameters than the CJS Model. Individual points on the graphic can be selected to identify the period and population they represent. To select a point, move the mouse cursor to that point, then left-click. The point that is displayed above **as** a small white circle has been selected. The description of the selected point appears in the footer of the window. Here, the selected point is from population 5, period 3. The CJS capture estimate is 0.50, and the SURPH capture estimate is 0.84.

THIS PAGE INTENTIONALLY LEFT BLANK

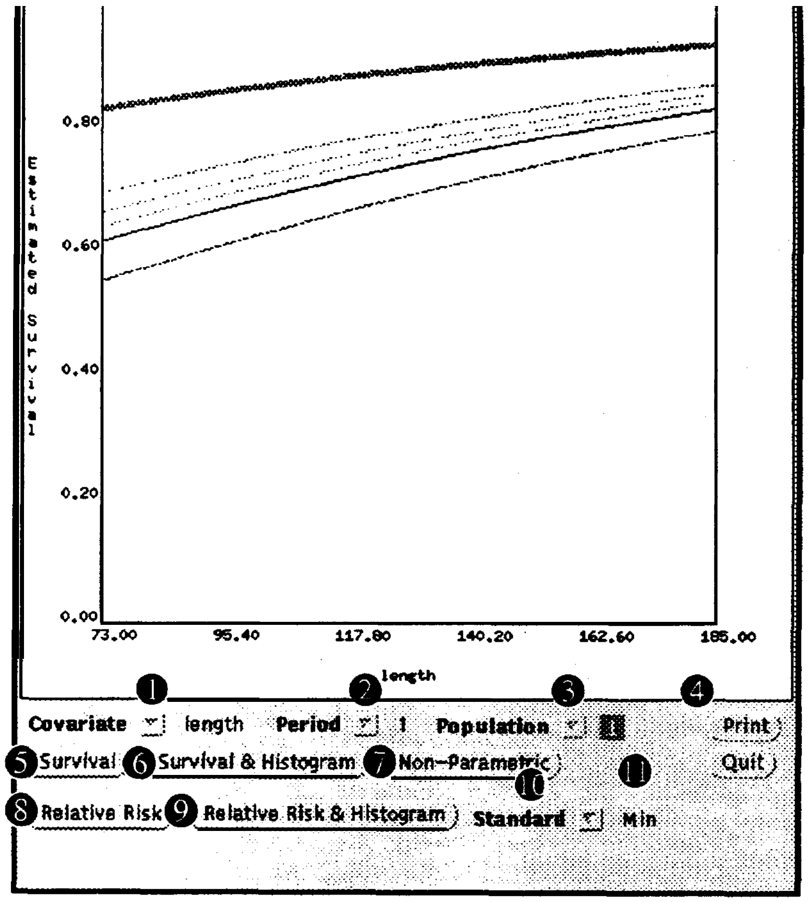


This window **(Models -> Survival Curves -> Group)** provides a graphical display of the estimated SURPH survival curve for a group covariate. The large circular points on the plot represent estimated survival probabilities from the **Cormack/** Jolly-Seber (Cormack 1964, **Jolly** 1965, Seber **1965)** model plotted against the group covariate value for each population. The colored line (dashed above) is the **fitted** survival model as a function of the current group covariate. This graphic, therefore, can be used in conjunction with the SURPH vs CJS graphic to determine how well the **SURPH** model fits the data. There are two pull-down menus available on the control panel at the bottom of the window:

❶ **Period Pull-Down Menu** - Determines the sampling period for the diplayed graphic.

❷ **Covariate Pull-Down** Menu - Determines the group covariate for the displayed graphic.

THIS PAGE INTENTIONALLY LEFT BLANK



This window **(Models -> Survival Curves -> Individual)** provides graphical displays of the fitted model of survival as a function of individual covariates. Several different graphics can be displayed on the same window. The graphic to be displayed depends upon the combination of choices you make from the control panel at the bottom of the window. The default display, shown when the window is **first** displayed, and illustrated above, shows the fitted survival curves for all populations. The covariate value is on the horizontal axis, and the estimated survival probability is on the vertical axis. The curve for each population has a distinct color, and the curve for the population currently selected on the Population Pull-Down Menu (see below) is double the thickness of those for the non-selected populations.

User options on the control panel are:

0 Covariate of Interest - This pull-down menu allows you to select the individual covariate to be displayed. Right-click to access the **pulldown** menu, then left-click on the desired covariate.

2 Period of Interest - This pull-down menu allows you to select the period to be displayed. Right-click to access the pull-down menu, then left-click on the desired period.

3 Population of Interest - This pull-down menu allows you to select the population to be displayed. Right-click to access the pull-down menu, then left-click on the desired population.

4 **Print -** This button allows you to print the graphic. Left-click on this button to send a copy of the current display to the printer.

5 **Survival -** Left-click to display fitted survival curves for all populations for the selected covariate and period (default, shown above).

6 **Survival & Hist -** Left-click to display fitted survival curves for the selected population, superimposed on the histogram of the selected covariate for the period of interest.
See Models: Survival Graphics: Individual Covariate (B) **for** an example of this display.

7 **Non-Parametric -** Left-click to display the fitted survival curve for the population, period and covariate of interest, along with a nonparametric **analogue.**
See Models: Survival Graphics: Individual Covariate (C) **for** an example of this display.

8 **Relative Risk -** Left-click to display the fitted relative risk curves for all populations for the selected covariate and **period.**
See Models: Survival Graphics: Individual Covariate (D) **for** an example of this display.

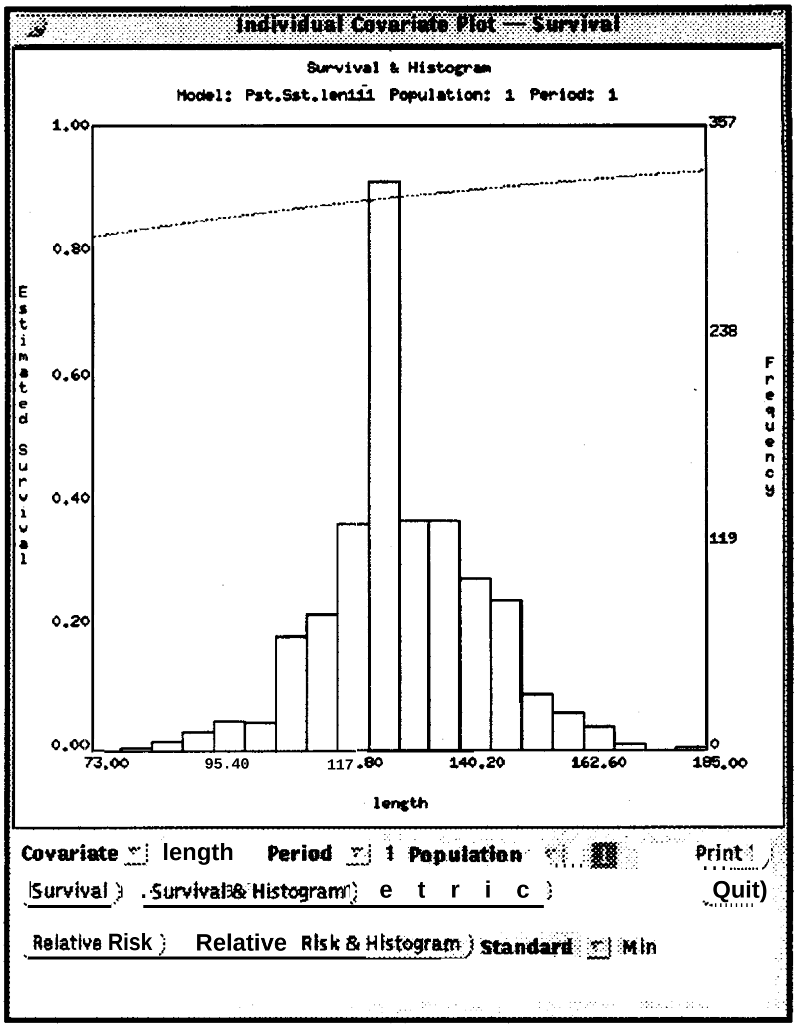
9 **Relative Risk & Hist -** Left-click to display the fitted relative risk curve for the selected population, superimposed on a histogram of the selected covariate for the **period** of interest.

See Models: Survival Graphics: Individual Covariate (E) **for** an example of this display.

10 Standard - This pull-down menu allows you to specify the standard covariate value for calculation of the relative risk. you can specify the risk as relative to either the minimum covariate value, the mean covariate value, or the maximum covariate value. Right-click to access the pull-down menu, then left-click on the desired standard.

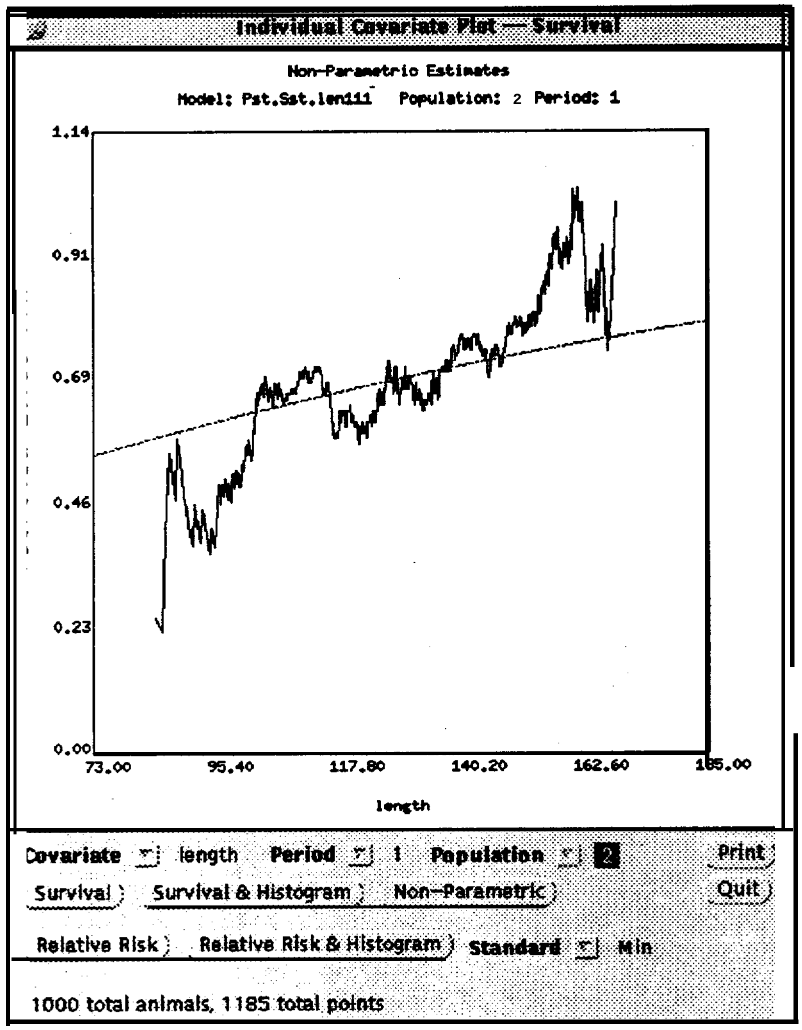
11 Quit - Left-click to dismiss this window.

THIS PAGE INTENTIONALLY LEFT BLANK



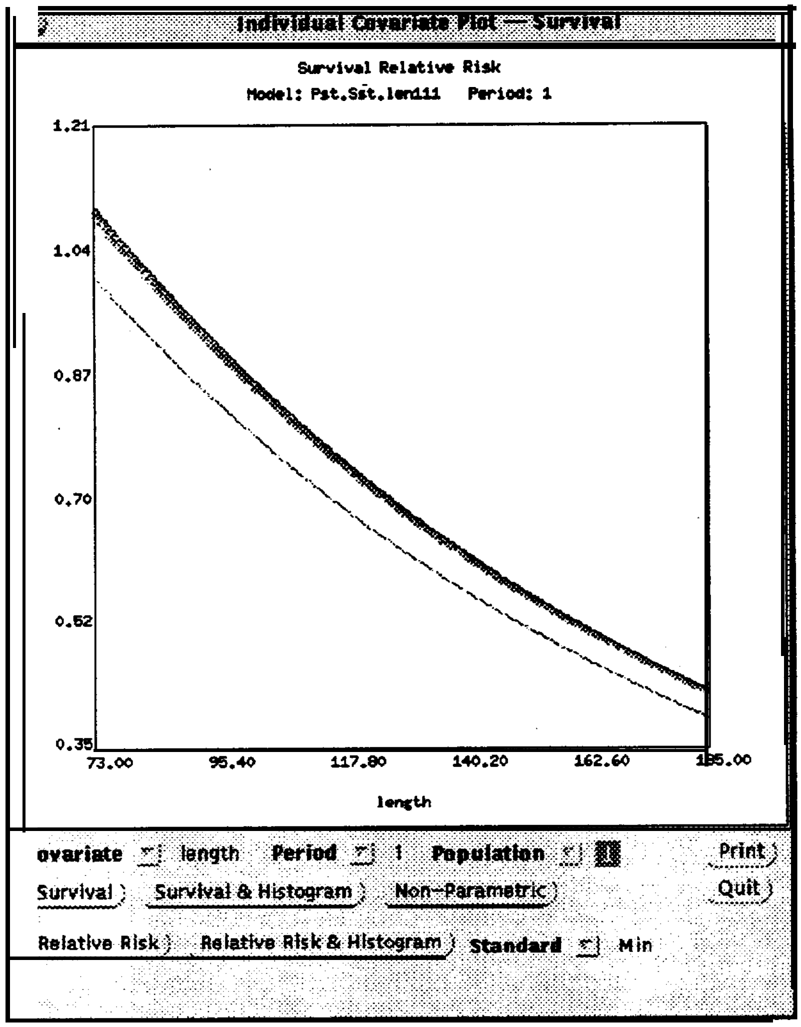
This display **(Models -> Survival Curves -> Individual)** shows the fitted survival curve for the population, covariate, and period of interest, superimposed over a histogram of the covariate for animals known alive at the start of the period displayed. While the "Survival and Histogram" display is active, you have the option of changing the population, covariate, or period of interest using the respective pull-down menus.

THIS PAGE INTENTIONALLY LEFT BLANK



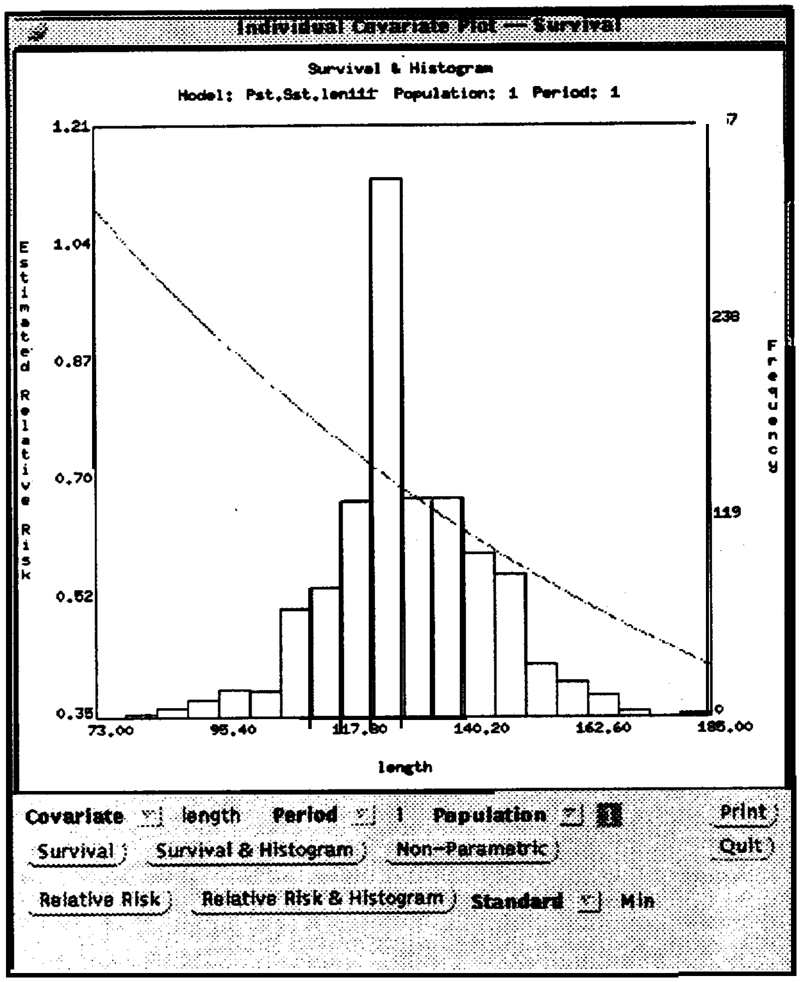
This display **(Models -> Survival Curves -> Individual)** shows the **fitted** survival curve (smooth) for the population, covariate, and period of interest, superimposed on a nonparametric analogue curve (jagged). For **further information** concerning the nonparametric curve, see Section 4.6.4 in the SURPH.l manual. While the **"Non-**Parametric" display is active, you have the option of changing the population, covariate or period of interest using the respective pull-down menus.

THIS PAGE INTENTIONALLY LEFT BLANK



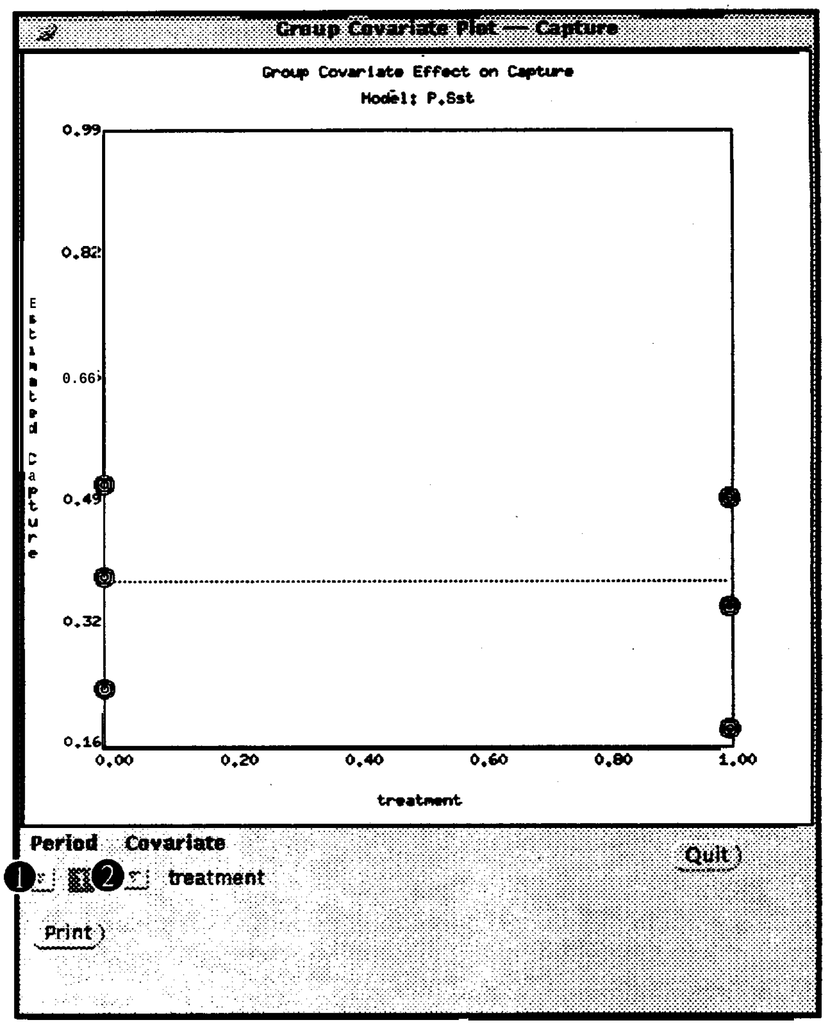
This display **(Models -> Survival Curves -> Individual)** shows the fitted relative risk curves for the covariate and period of interest for all populations. The "standard" covariate value for the risk can be changed so that the risk to the animal with the minimum value of the covariate is set to **"1",** the risk to the animal with the mean value of the covariate is set to **"1",** or the risk to the animal with the maximum value of the covariate is set to "1". Because the risk is given in relative terms, changing the standard affects the scale of the curve, but does not affect the shape of the curve, nor the conclusions to be drawn. While the "Relative Risk" display is active, you have the option of changing the population, covariate, or period of interest using the respective pull-down menus.

THIS PAGE INTENTIONALLY LEFT BLANK



This display **(Models -> Survival Curves -> Individual)** shows the fitted relative risk curve for the population, covariate, and period of interest, superimposed over a histogram of the individual covariate of interest for the population of animals released during the period specified. The "standard" covariate value for the risk can be changed so that the risk for the animal with the minimum **value** of the covariate is set to 1.0, the risk for the animal with the mean value of the covariate is set to 1.0, or the risk for the animal with the maximum value of the covariate is set to 1.0. Because the risk is given in relative terms, changing the standard affects the scale, but does not affect the shape of the curve, nor the conclusions to be drawn. While the "Survival and Histogram" display is active, you have the option of changing the population, covariate, or period of interest using the respective pull-down menus.

THIS PAGE INTENTIONALLY LEFT BLANK

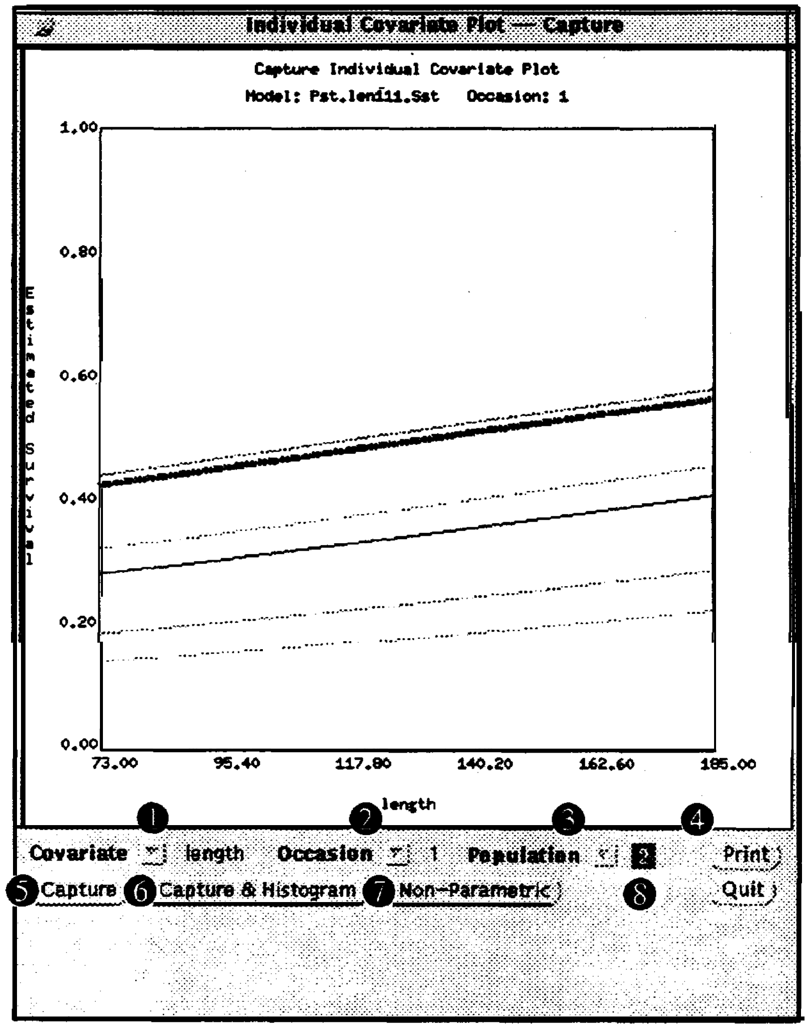


This window **(Models -> Capture** Curves **-> Group)** provides a graphical display of the estimated **SURPH** capture curve for a group covariate. The large circular **points** on the plot represent estimated capture probabilities from the **Cormack/Jolly-Seber** (Cormack 1964, Jolly 1965, Seber 1965) model plotted against the group covariate value for each population. The colored line (dashed above) is the fitted captured model as a function of the current group covariate. This graphic, therefore, can be used in conjunction with the **SURPH** vs CJS graphic to determine how well the SURPH model fits the data. There are two pull-down menus available on the control panel at the bottom of the window:

❶ Occasion Pull-Down Menu - Determines the sampling occasion for the diplayed graphic.

❷ Covariate Pull-Down Menu - Determines the group covariate for the displayed **graphic.**

THIS PAGE INTENTIONALLY LEFT BLANK



This window **(Models -> Capture Curves -> Individual)** provides graphical displays of the fitted model of capture as a function of individual covariates. Several different graphics can be displayed on the same window. The graphic to be displayed depends upon the combination of choices you make from the control panel at the bottom of the window. The default display, shown when the window is first displayed and illustrated above, shows the fitted capture curves for all populations. The covariate value is on the horizontal axis, and the estimated capture probability is on the **vertical axis**. The curve for each population has a distinct color, and the curve for the population currently selected on the Population Pull-Down Menu (see below) is double the thickness of those for the non-selected populations.

User options on the control panel are:

0 Covariate of Interest - This pull-down menu allows you to select the individual covariate to be displayed. Right-click to access the **pulldown** menu, then left-click on the desired covariate.

2 Occasion of Interest - This pull-down menu allows you to select the occasion to be displayed. Right-click to access the pull-down menu, then left-click on the desired occasion.

3 Population of Interest - This **pull-down** menu allows you to select the population to be displayed. Right-click to access the pull-down menu, then left-click on the desired population.

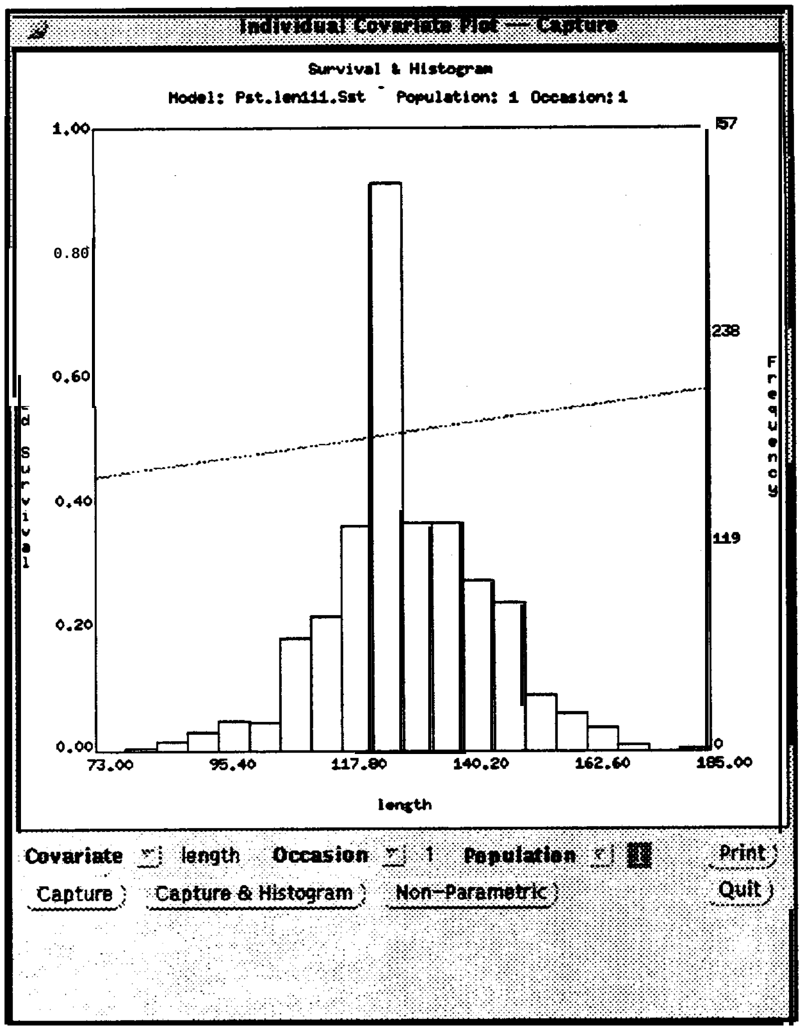
4 **Print -** This button allows you to print the graphic. Left-click on this button to send a copy of the current display to the printer.

5 **Capture -** Left-click to display fitted capture curves for all populations for the selected covariate and period (default, shown above).

6 **Capture & Histogram -** Left-click to display fitted capture curves for the selected population, superimposed on the histogram of the selected covariate for the **period** of interest.
See Models: Capture Graphics: Individual Covariate (B) **for an** example of this display.

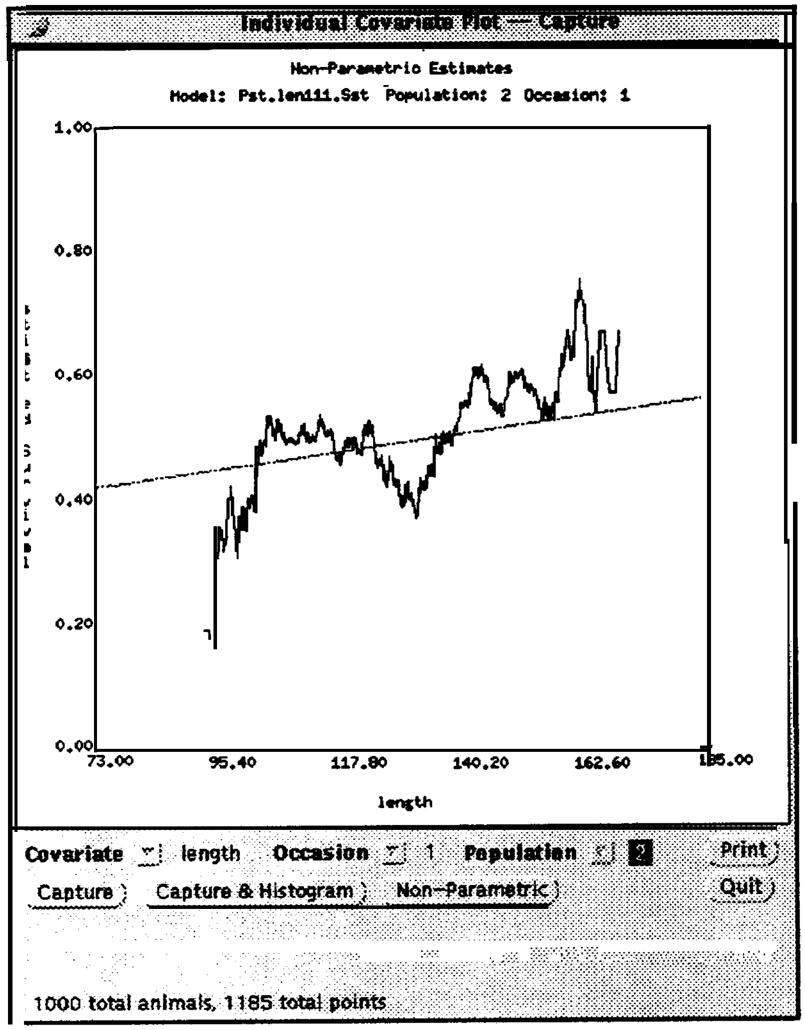
7 **Non-Parametric -** Left-click to display the fitted capture curve for the population, period and covariate of interest, along with a nonparametric **analogue.**
See Models: Capture Graphics: Individual Covariate (C) **for an** example of this display.

8 **Quit -** Left-click to dismiss this window.



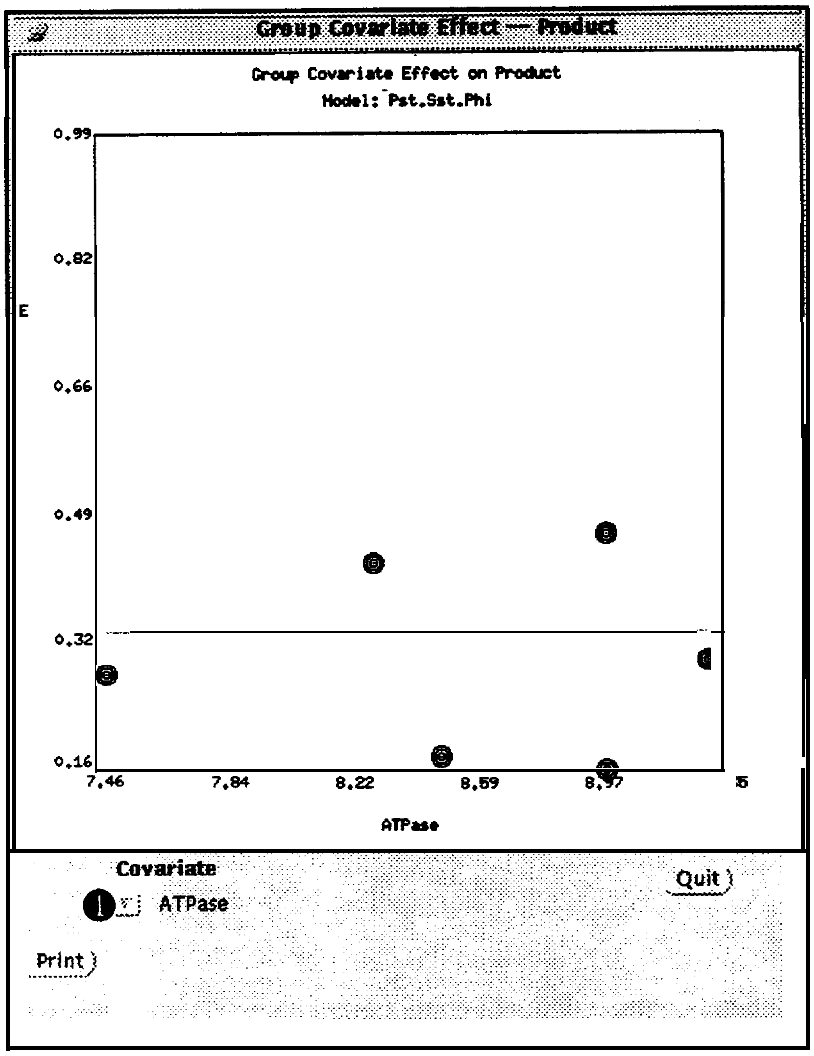
This display **(Models -> Capture Curves -> Individual)** shows the fitted capture curve for the population, covariate, and occasion of interest, superimposed over a histogram of the covariate for animals known alive at the time of the selected sampling occasion. While the "Capture and Histogram" display is active, you have the option of changing the population, covariate, or occasion of interest using the respective pull-down menus.

THIS PAGE INTENTIONALLY LEFT BLANK



This display **(Models -> Capture Curves -> Individual)** shows the fitted capture curve (smooth) for the population, covariate, and occasion of interest, superimposed on a nonparametric analogue curve (jagged). For further information concerning the nonparametric curve, see Section 4.6.2.5 of the SURPH.l manual. While the **"Non-**Parametric" display is active, you have the option of changing the population, covariate or occasion of interest using the respective pull-down menus.

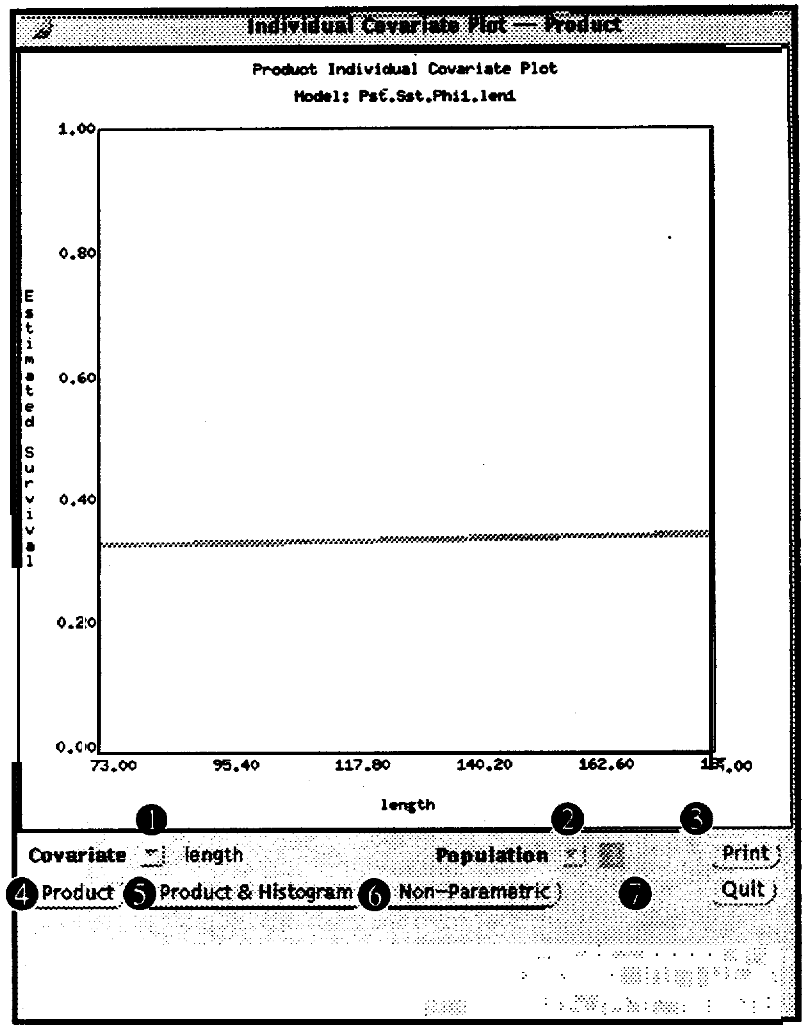
THIS PAGE INTENTIONALLY LEFT BLANK



This window **(Models -> Product Curves -> Group)** provides a graphical display of the estimated **SURPH** product curve for a group covariate. The large circular points on the plot represent estimated product probabilities from the **Cormack/Jolly-Seber** (Cormack 1964, Jolly 1965, Seber 1965) model plotted against the group covariate value for each population. The colored line (dashed above) is the fitted survival model as a function of the current group covariate. This graphic, therefore, can be used in conjunction with the **SURPH** vs CJS graphic to determine how well the SURPH model fits the data. There are two pull-down menus available on the control panel at the bottom of the window:

0 Covariate Pull-Down Menu - Determines the group covariate for the displayed graphic.

THIS PAGE INTENTIONALLY LEFT BLANK



This window **(Models -> Product Curves -> Individual)** provides graphical displays of the **fitted** model of product as a function of individual covariates. Several different graphics can be displayed on the same window. The graphic to be displayed depends upon the combination of choices you make from the control panel at the bottom of the window. The default display, shown when the window is first displayed, and illustrated above, shows the fitted product curves for all populations. The **covariate** value is on the horizontal axis, and the estimated product probability is on the vertical axis. The curve for each population has a distinct color, and the curve for the population currently selected on the Population Pull-Down Menu (see below) is double the thickness of those for the non-selected populations.

User options on the control panel are:

1. Covariate of Interest - This pull-down menu allows you to select the individual covariate to be displayed. Right-click to access the pull-down menu, then left-click on the desired covariate.

2. Population of Interest - This pull-down menu allows you to select the population to be displayed. Right-click to access the pull-down menu, then left-click on the desired population.

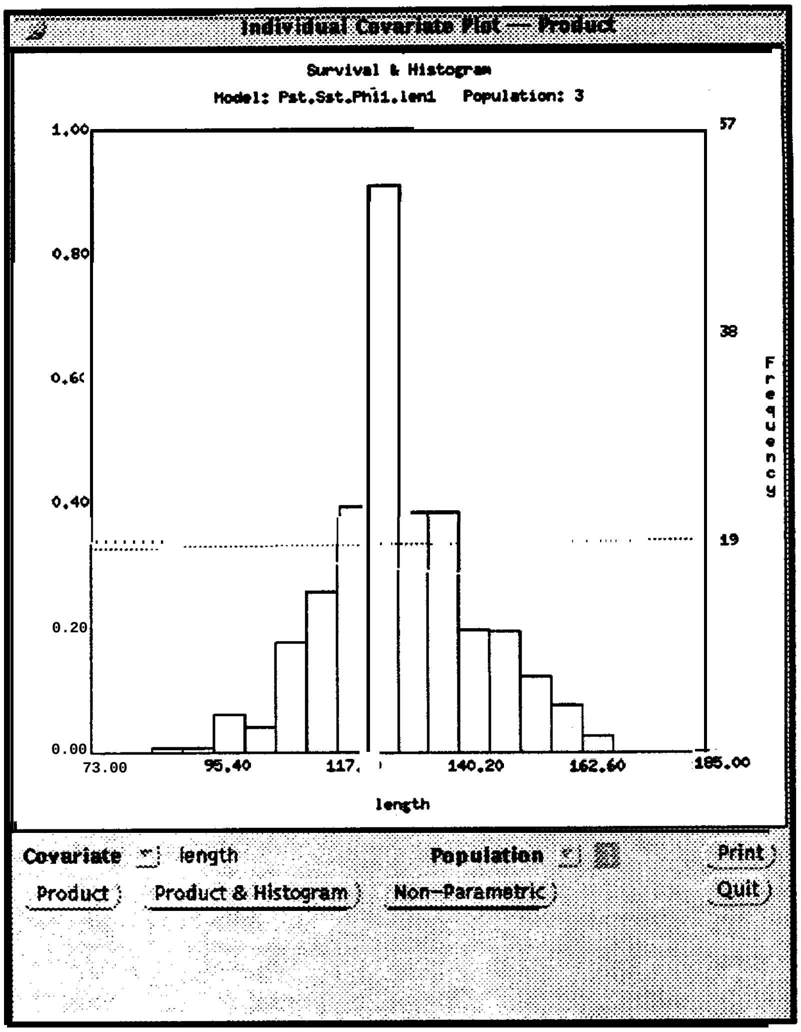
3. **Print -** This button allows you to print the graphic. Left-click on this button to send a copy of the current display to the printer.

4. **Product -** Left-click to display fitted product curves for all populations. for the selected covariate (default, shown above).

5. **Product & Histogram -** Left-click to display fitted product curve for the selected population, superimposed on the histogram of the selected covariate of interest.
   See Models: Product Graphics: Individual Covariate (B) **for an** example of this display.

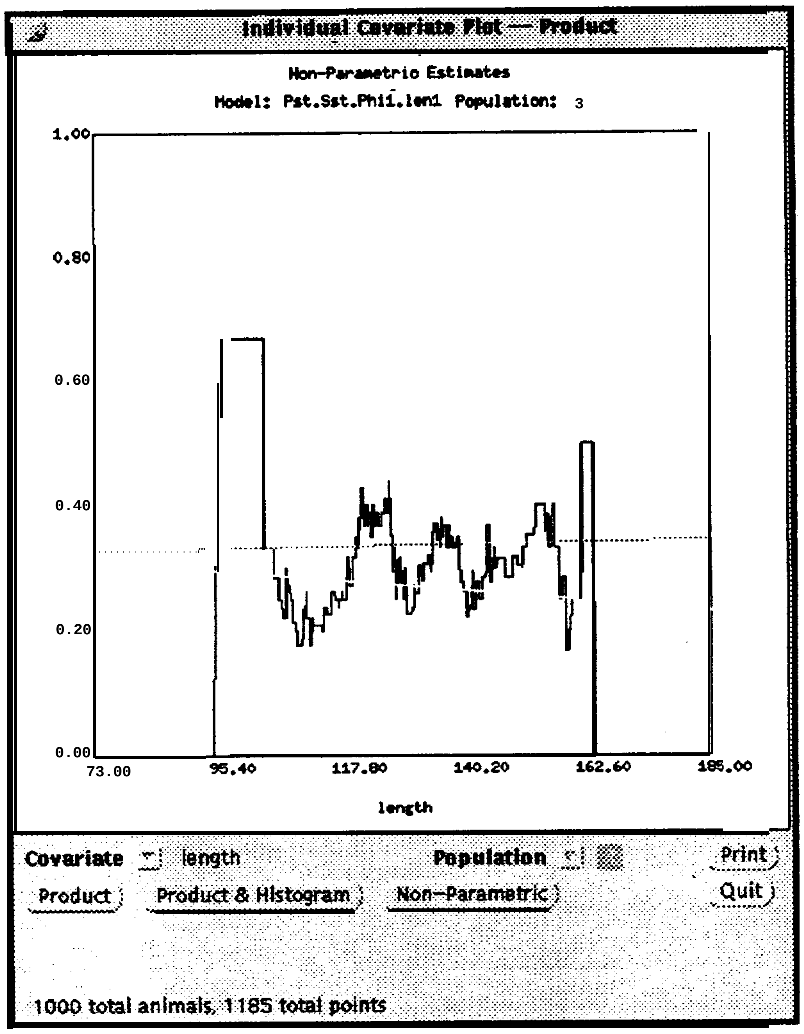
6. **Non-Parametric -** Left-click to display the fitted product curve for the population and covariate of interest, along with a nonparametric analogue.
   See Models: Product Graphics: Individual Covariate (C) for an example of this display.

7. **Quit -** Left-click to dismiss this window.



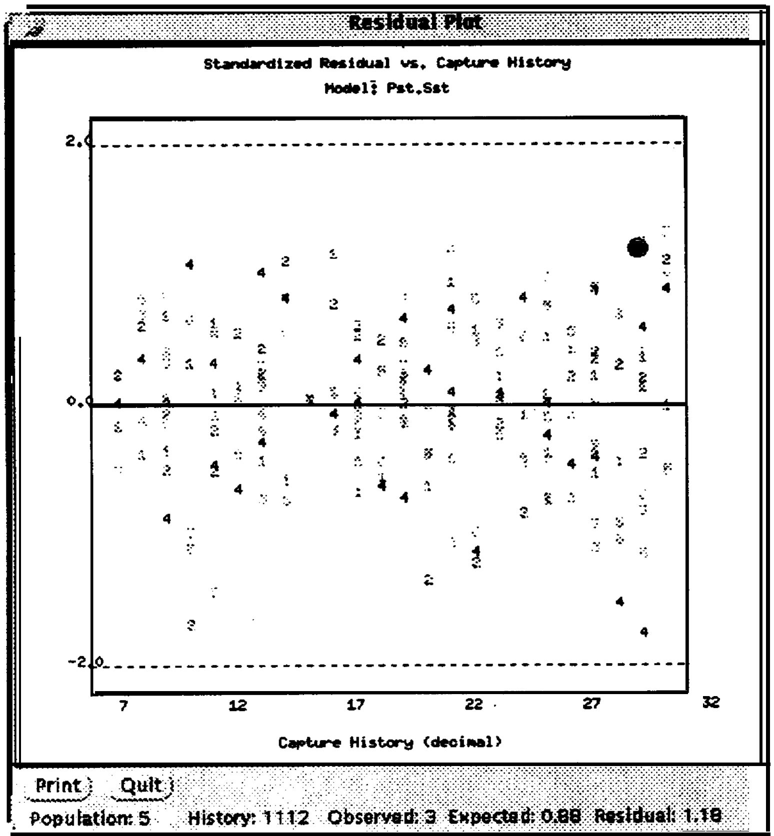
This display **(Models -> Product Curves -> Individual)** shows the fitted product curve for the population and **covariate** of interest, superimposed over a histogram of the covariate for animals known alive during the final period. While the "Product and Histogram" display is active, you have the option of changing the population or covariate of interest using the respective pull-down menus.

THIS PAGE INTENTIONALLY LEFT BLANK



This display **(Models -> Product Curves -> Individual)** shows the fitted product curve (smooth) for the population and covariate of interest, superimposed on a nonparametric analogue curve (jagged). For further information concerning the nonparametric curve, see Section 4.6.2.5 of the SURPH.l manual. While the **"Non-**Parametric" display is active, you have the option of changing the population or covariate of interest using the respective pull-down menus.

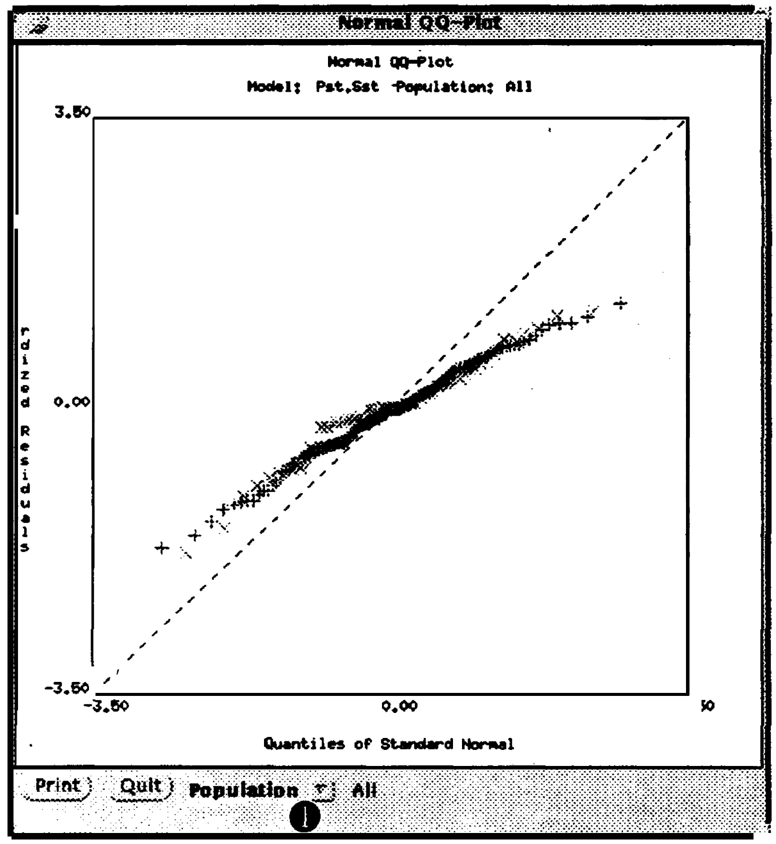
THIS PAGE INTENTIONALLY LEFT BLANK



This plot **(Models -> Residual Plots -> vs. Capture Histories)** displays the Anscombe standardized residuals (Anscombe standardized residuals are distributed **Assymptotically** Normal **(0,1); McCullagh** and Nelder 1983) versus the capture histories. Buttons in a control panel at the bottom **allow you to Print** a copy of this window or to **Quit** (dismiss) this window. The left mouse button can be used to select individual points on the plot (indicated by the small white circle). When a point is selected, the capture history, the expected value of the capture history and the observed value of the capture history are displayed for the selected point in the window footer. The plotting characters of the points on the plot designate the population from which the residual was calculated. Thus, for each possible capture history, there are p points on the plot, where p is the the number of populations in the data set.

Recall that capture histories resemble binary numbers (e.g., 10000, = 16_{10}). The x-axis, therefore, displays capture histories converted into decimal equivalents.

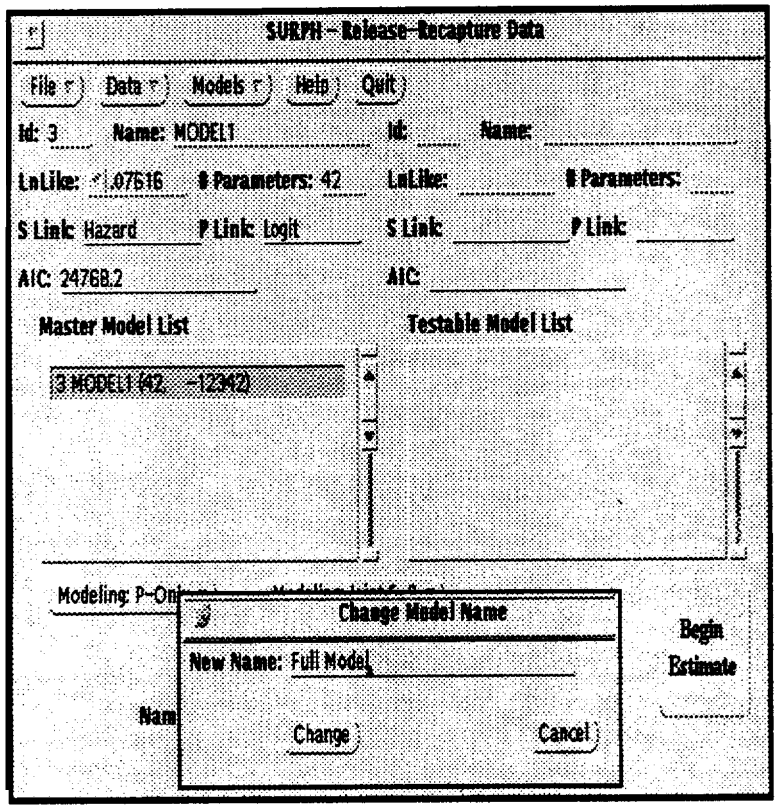
THIS PAGE INTENTIONALLY LEFT BLANK



This plot **(Models -> Residual Plots -> QQ-Plot)** displays the **Anscombe** standardized residuals (McCullagh and Nelder 1983) for the model currently selected in **the Master Model List** versus their expected value under the Normal distribution (Zar 1984). Purple crosses indicate capture histories of removed animals, whereas green x's indicate complete capture histories of tagged animals. Buttons on the control panel at the bottom of this plot **allow you** to **Print** a **copy** of **this** window, or **Quit** (dismiss) this window.

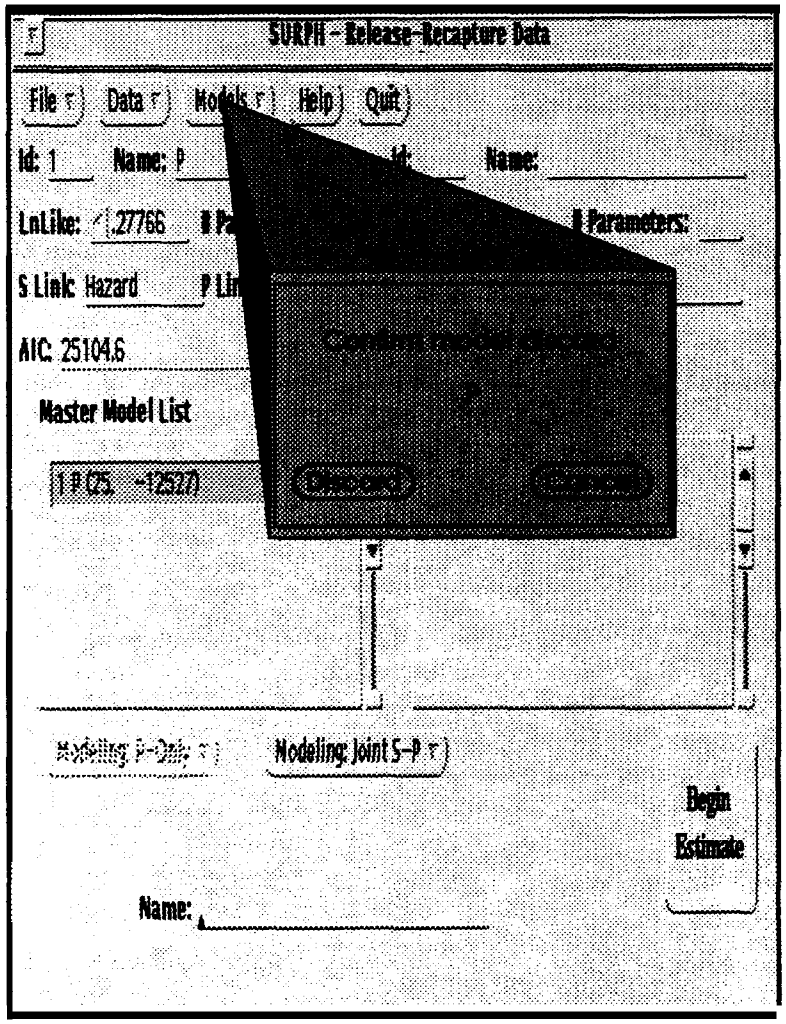
❶ Population - This Pull Down Menu allows you to specify the population from which the residuals will be drawn. You may specify either residuals for an individual population or the residuals for all populations.

THIS PAGE INTENTIONALLY LEFT BLANK



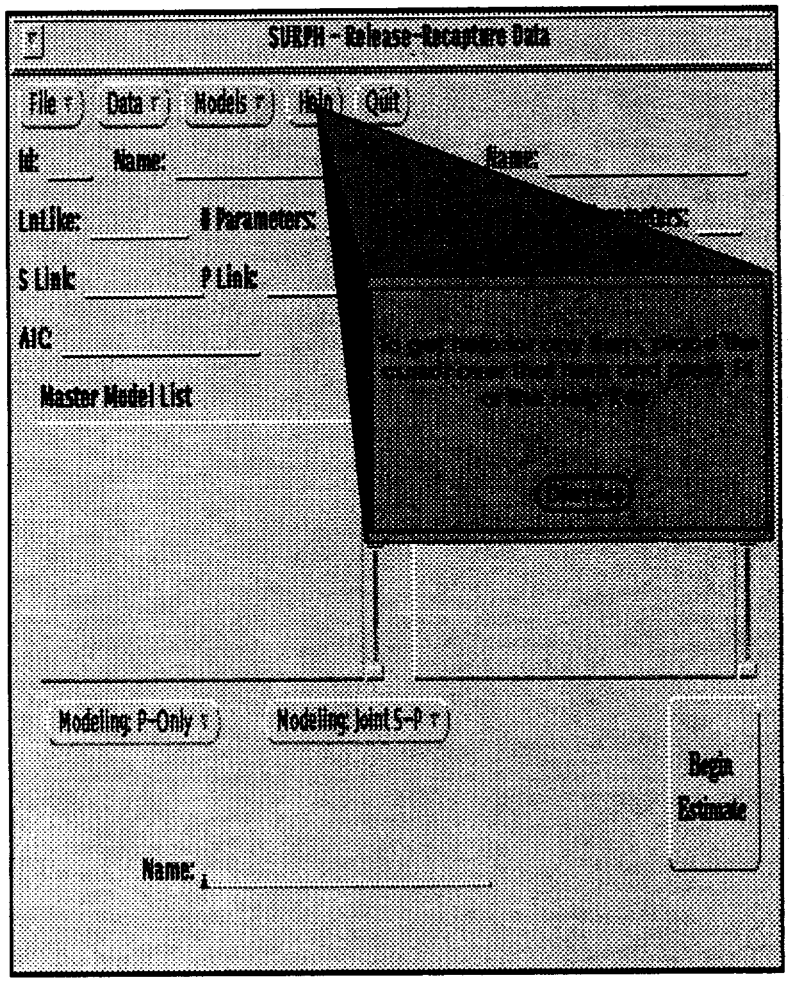
This small pop-up window **(Models -> Change Name) allows you** to change the name of the model currently selected within **the Master Model List**. **To** change the name of the selected model, left-click on the **New Name Field, type the** new name, then left-click on the **Change Button. In this** example, the model currently named *'MODEL 1" will be renamed "Full Model" upon left-clicking the **Change Button.**

THIS PAGE INTENTIONALLY LEFT BLANK



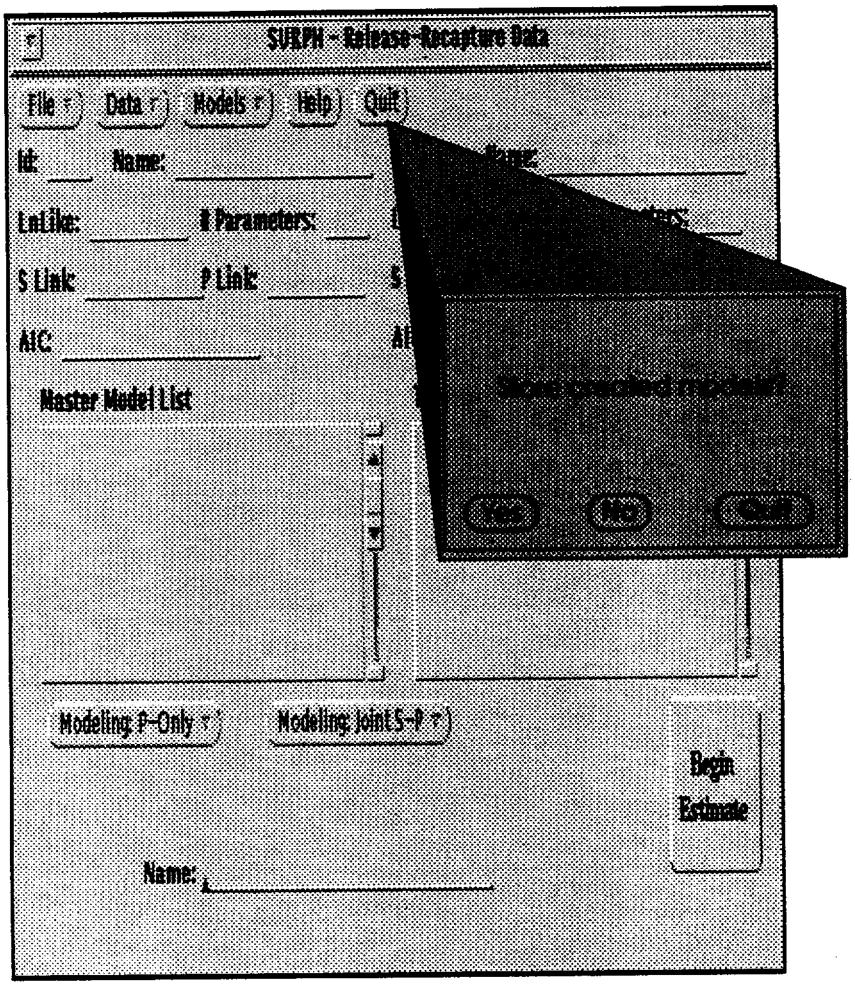
This dialog box **(Models -> Discard) allows you to** remove models that are no longer needed **in** SURPH. **In the Master Model List,** select the model you wish to delete by left-clicking on the model name. Then go to the Discard Models command **(Models -> Discard),** and depress the left-click on the **Discard** menu option. **The Cancel Button allows you** to exit this command without deleting a model. Once a model has been deleted, it is inaccessible, and cannot be undeleted. The only way you can "retrieve" a model that has been deleted is to estimate a new model that has the exact same parameterization as the one deleted.

THIS PAGE INTENTIONALLY LEFT BLANK



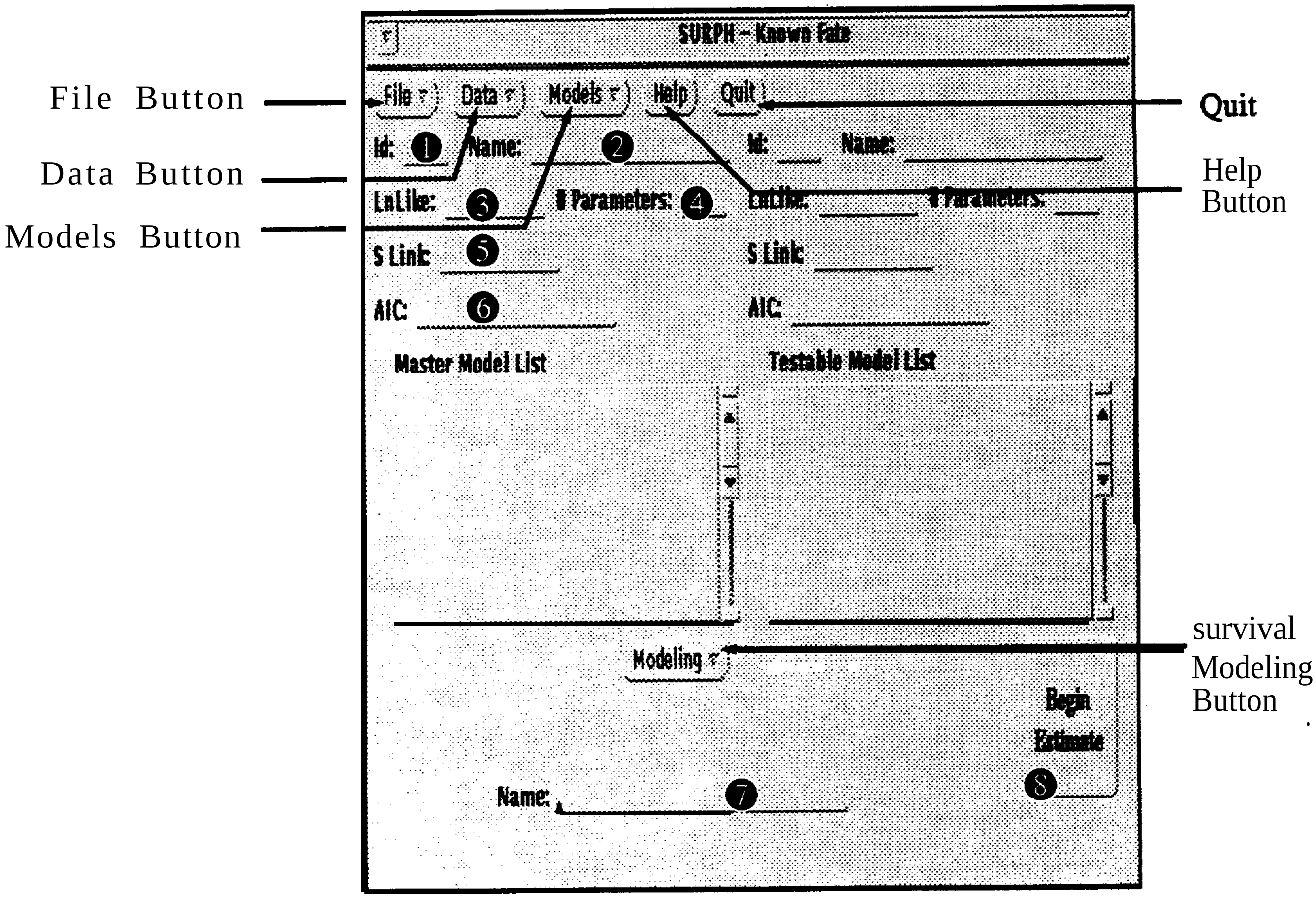
Left-clicking on the **Help Button** summons the Help Dialog Box. This dialog box window tells you how to get help for any item of the **SURPH** program.

THIS PAGE INTENTIONALLY LEFT BLANK



Left-clicking on the **Quit Button** summons the Quit Dialog Box. If models have been created since the most recent Store Models command, you are asked if the models should be saved before quitting the program. To choose an answer, left-click on the answer desired. If you answer "Yes", a Store Model Pop-Up window appears (see File: Store Models). You then enter the name of the file in which the models are to be stored. Once the models have been stored, the SURPH session ends. If you answer "No", the SURPH session is ended, without storing the estimated models. If you answer **"Cancel", the dialog** box is dismissed and the SURPH session continues.

THIS PAGE INTENTIONALLY LEFT BLANK



This is the SURPH Base **Window** for Known-Fate data files, the first window that appears after you select the Known-Fate modeling option from the SURPH Introductory Window. The SURPH Introductory Window appears when the program is started with a command line:

> surph ***datajile***

All data summaries, modeling operations, graphics, and diagnostics are accessed from the SURPH Base Window.

Overview of Command Buttons

The pull-down menu **accessed** by **the File Button is used** to load data files, and to load or store model files. **The File** pull-down can also be used to end the SURPH session. Because the data file was specified on the command line, the Load Data command **(File -> Load Data)** is disabled.

The pull-down menu accessed by the **Data Button** has choices for computed Binomial probability estimates and Sampling Summary Statistics; graphical data display; data transformations; and echoing input data to the screen.

The pull-down menu **accessed** by **the Models Button allows** display of model estimates and graphical displays of model fit. This button also allows you to discard models and change model names.

The Help Button explains the use of the various buttons. To see further instructions, left-click on this button.

The Quit Button ends the **SURPH** session.

The Modeling Button is used to specify the **parameterization** of survival models. These buttons have pull-down menus that allow you to model the relationships between factors measured at the population and individual level, and the survival processes. Menu choices under **the Modeling Button** allow you to model the survival processes at the population or individual level, and allow you to specify the link function. For further descriptions of the functions associated with this button, see

Survival Modeling: Pull-Down Menu

The SURPH Known-Fate Base Window has two lists of estimated models, the **Master Model List** and the **Testable Model List. The Master Model List** contains all models that have been estimated during the current SURPH session, and those that have been loaded from previous SURPH sessions. The Testable Model List contains models that can be compared to the model that has been selected from the **Master Model List.** For a model to appear within the **Testable Model List,** either the model within the **Testable Model List is** nested within the model selected from the **Master Model List,** or the model selected from the **Master Model List** is nested within **the** model **in the Testable Model List** **(see** Testing Nested Models: Master and Testable Lists **).**

Numbered Features

The numbered spaces 1-7 refer to attributes of the model selected from the **Master Model List.** Analogous spaces in the upper right-hand portion of the **SURPH** Base Window refer to attributes of the model selected **within the Testable Model List.**

0 Displays the Model ID Number of the selected model **in the Master Model List.**

❷ Displays the name of the selected model **in the Master Model List.**

❸ Displays the Log-Likelihood of the selected model **in the Master Model List.**

❹ Displays the number of parameters estimated **in the** selected model **in the Master Model List.**

❺ Displays the link function for the survival parameters of the selected model in the **Master Model List.**

❻ Displays Akaike's Information Criterion (Akaike 1973) **-** This is used to help determine the best fitting models when the models are not nested.

The space numbered 7, and the button numbered 8 are used to create new models:

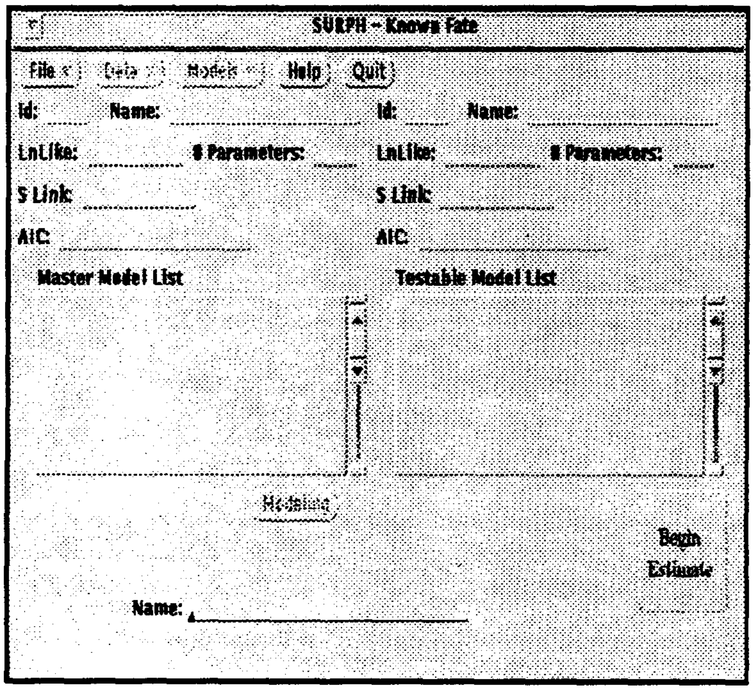
❼ This field allows you to name the model prior to running the estimation procedure. If you does not specify a name for the model, the name "Model 'n"' is assigned, where **"n"** is the Model ID Number. If you have a model you wish to rename, **see the** section for the **Models Button.**

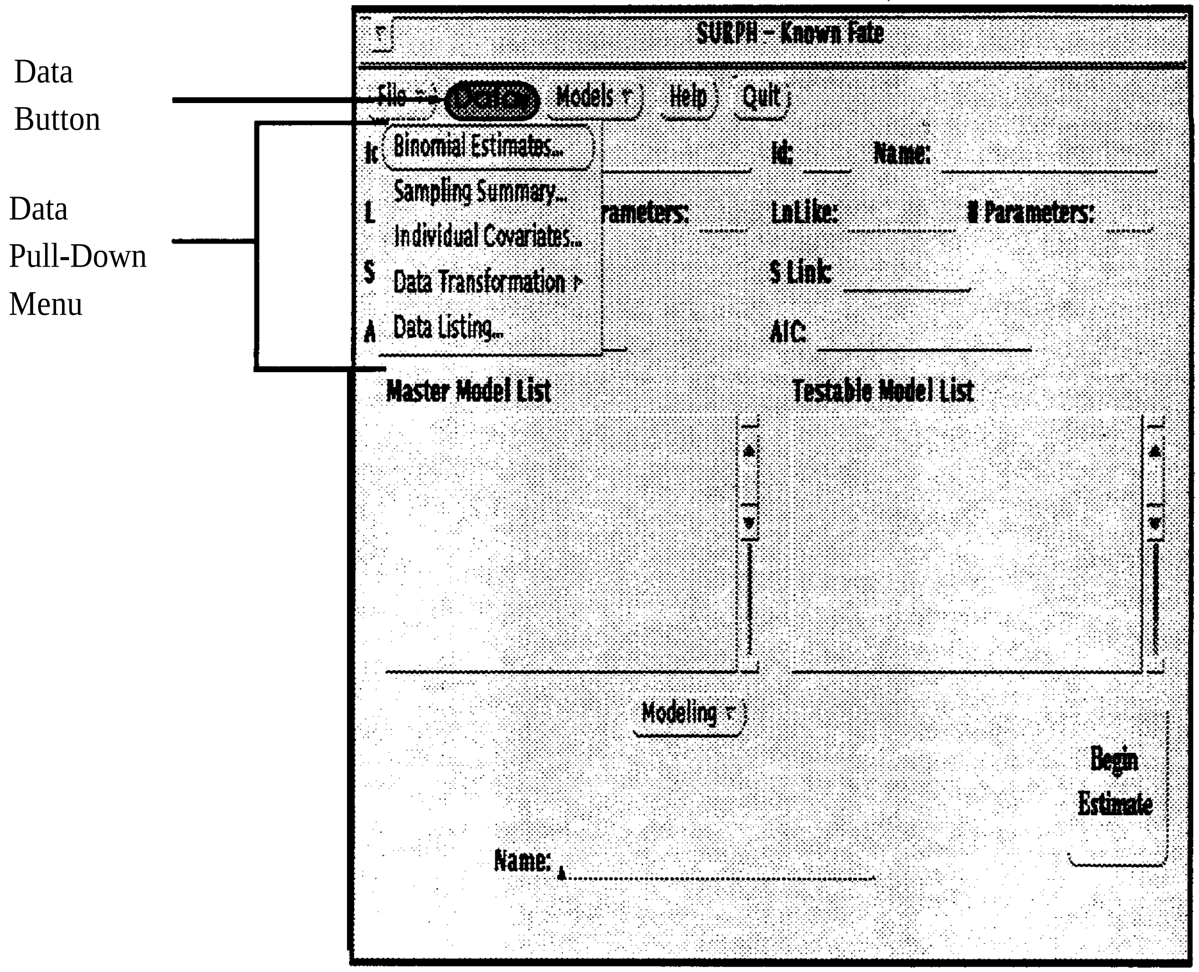
❽ Left-clicking on this button initiates the numerical optimization routine to estimate the parameters of the model **specified in the** modeling options **under Modeling.**

If you start **SURPH** without specifying an input data file on the command line:

> surph

not all buttons are activated (see below). Before these de-activated buttons can be used, data must be loaded. To load data, right-click on **the File Button. A** pull-down menu will appear. Left-click on **Load Data. The Load Input File** window will appear. If your data file is in the directory from which you started **SURPH,** left-click **in the File** text field and type the name of the file that contains your data. If your data is in a different directory, left-click **in the Directory** text field and change the **directory** path **in the Directory** field, then left-click in the **File** field and enter the file name in **the** blank **File** field. Left-click on the **Load Button in the Load Input Data** window to load the data.





To access this pull-down menu, right-click on **the Data Button.**

The commands available from this menu specify data summaries and transformations:

Binomial Estimates - Provides the binomial estimates of survival probabilities. For further information, see Data: Binomial Estimates .

Sampling Summary - Provides a summary of the sampling results for each population during each period. For further information, see Data: Sampling Summary .

Individual Covariates - Displays the graphical representation (histograms or cumulative distribution functions) of the data, by covariate, by population, by period. For further information, see Individual Covariates - Histograms or Individual Covariates - CDF Plot

Data Transformation - Allows transformations of covariate data. Examples of transformations are squares, square roots, cross products, reciprocal and natural log. For further information, see Data Transformations: Group Covariates or Data Transformations: Individual Covariates

Data Listing - Gives a complete listing of the input data in a formatted display. For further information, see Data Listing .

Survival Estimates (Binomial Model)

Print Quit

```
Survival Estimates From Binomial Model

Dataname:Simulated telemetry data from telem_sim_win
 (file: /home/ruddy/surph/public/telem.dat)

                Period
Population |     1              2              3              4
-----------+--------------------------------------------------------
         1 | 0.878  43149 ) 0.571  40/70 ) 0.462  18/39 ) 0,611 ( 11/18 )
         2 | 0.681  32147 ) 0.627  37/59 ) 0.667  24136 ) 0.458 ( 11/24 )
         3 | 0.673  33149 ) 0.542  32/59 ) 0.484  15/31 ) 0.750 (  9112 )
         4 | 0.646  31/48 ) 0,491  28/57 ) 0.440  11125 ) 0,444 (  4/9  )
         5 | 0.660  31/47 ) 0,579  33/57 ) 0,438  14/32 ) 0,429 (  6114 )
         6 | 0.532  25/47 ) 0.423  22/52 ) 0.250   5/20 ) 0,750 (  3/4  )
```

This text window **(Data -> Binomial Estimates)** shows the binomial survival probability estimates. The number in parentheses beside each probability estimate show the fraction of the number of animals that were released and that were subsequently detected (the numerator) to the total number of animals released (the denominator) (i.e., For Population 1 in Period 1, the estimated binomial survival probability is 0.878, which is equal to 43/49; 43 of the 49 animals released on the first occasion in the fist population were subsequently detected).

THIS PAGE INTENTIONALLY LEFT BLANK

Sample Summary

Print Quit

Release and Sampling Summary

Datanare: Simulated telemetry data from telem_sim_win
(file: /mounts/ruddy/warren/public/telem.dat)

Population 1	Number Marked	Survived to Occasion 2	3	4	5
Marking Occasion 2	28	42:1	23:1	11:0	6:0
3	0		16:0	7:0	4:1
				0:0	0:0
4	0				0:0

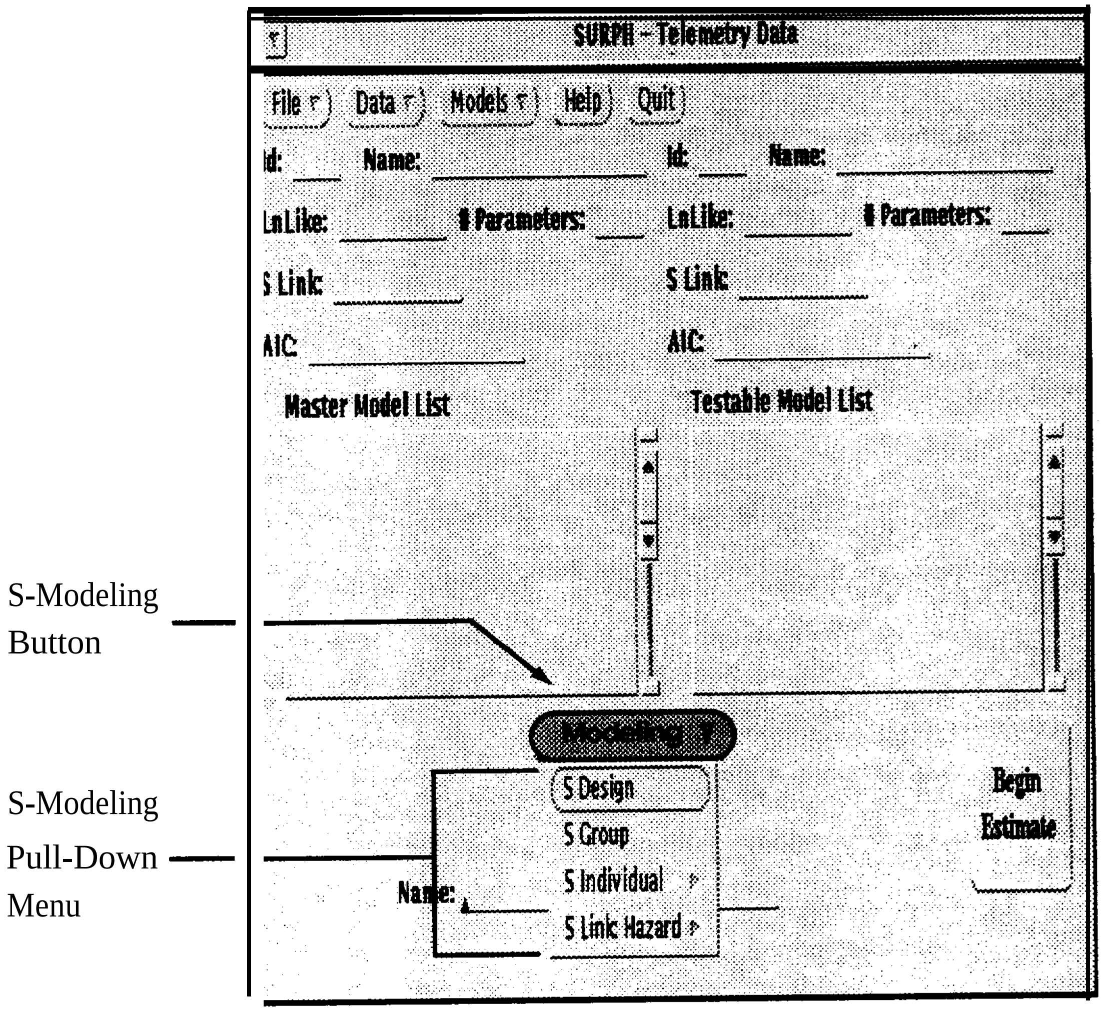
Population 2	Number Marked	Survived to Occasion 2	3	4	5
Marking Occasion 1	47	29:3	19:o	14:0	7:0
2	30		17:1	10:0	3:1
3	0			0:0	0:0
4	0				0:0

Population 3	Number Marked	Survived to Occasion 2	3	4	5
ttarklng Occasion 1	49	31:2	17:o	7:3	5:0
2	28		14:1	5:0	4:0
3	0			0:0	0:0
4	0				0:0

This text pane **(Data -> Sampling Summary)** display numbers of animals released and subsequently recaptured. There is one table for **each** population. The elements of the table for each population are:

1 Marking Occasion	Occasion when the animal was marked and released.
2 Number Marked	Number of animals marked and released at this occasion. This total only includes animals released for the fist time.
3 **Survived to** Occasion	For each marking occasion, the number of animals that survived until each subsequent occasion. The number to the left of the colon is the total number of animals known to have survived and remained at risk in the next period. To the right of the colon is the number "removed" from the sample for occasion **"i",** either because they were known to be dead on occasion "i", or because their radio-tags failed before the next **samplng** occasion. For example, 47 **animals** were marked on occasion 1 in population 2. Of these, 29 were detected alive on occasion 2 and were known to have working transmitters. Three animals detected on **oocasion** 2 were removed from the sample.

THIS PAGE INTENTIONALLY LEFT BLANK



This is the pull-down menu displayed when you right-click on **the Modeling Button.**

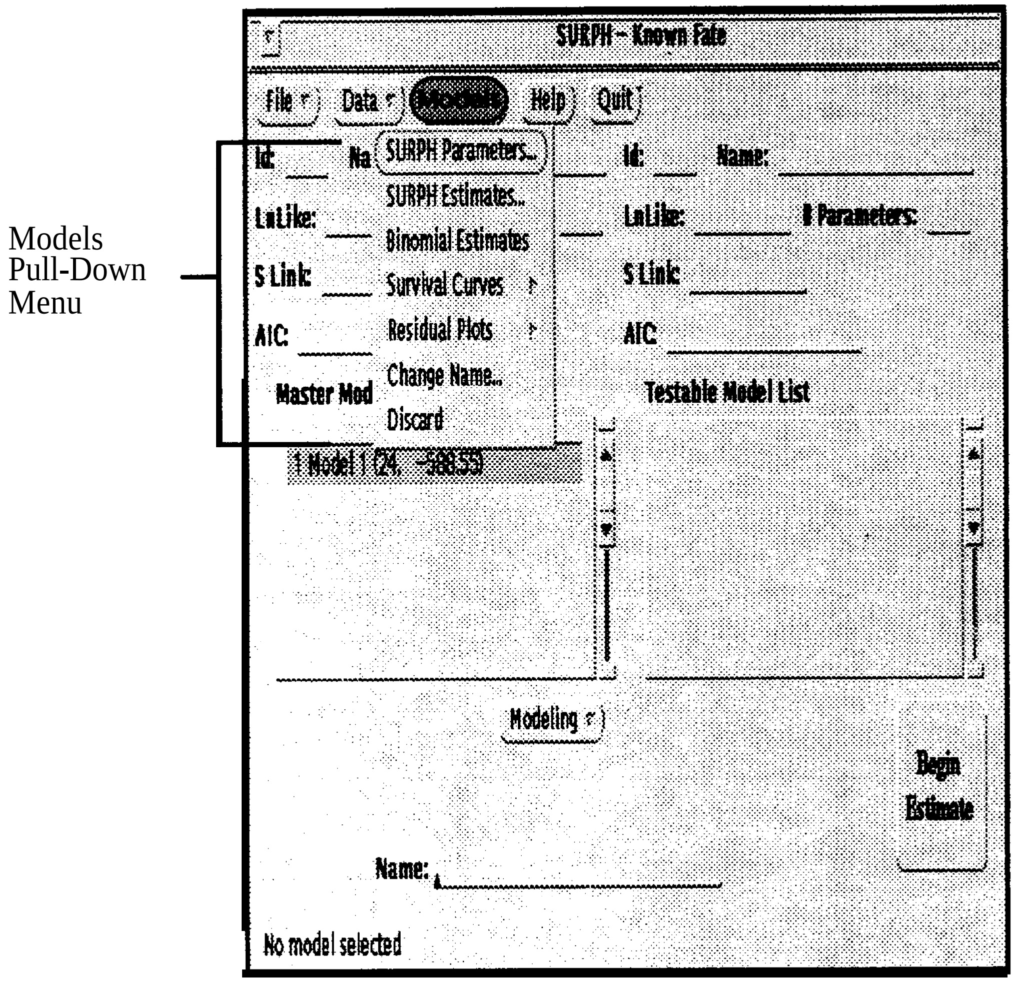
The commands available from this submenu specify models for survival probabilities:

S Design - Specify effects of the experimental design factor on survival probabilities. A Button Pad will appear that allows you to specify period and population-specific effects **(see** Survival Modeling: Design Pop-Up **).**

S **Group** - Specify group covariate effects on survival probabilities. A Button Pad will appear that allows you to specify parameters for period-specific group covariate effects (see Survival Modeling: Group Covariates).

S **Individual** - Specify individual covariate effects on survival probabilities. For each covariate selected, a Button Pad will appear. This Button Pad allows you to specify parameters for period and population-specific effects (see Survival Modeling: Individual Covariates).

S **Link** - Specify the function that links covariate effects to survival probabilities. The default link is Hazard. Alternatively, the Logit link function can be specified (see Survival Modeling: Link Function).



This is the pull-down window **accessed** by right-clicking on the **Models Button** while analyzing known-fate data. The Model **Pull-Down** Menu has choices to display SURPH survival estimates; view graphics that illustrate group and individual covariate **effects** on survival probabilities; view residual plots to investigate goodness-of-fit; change model names; and discard models. All actions pertain to the model currently selected in **the Master Model List.**

SURPH Parameters - Displays **text** window with estimated parameter values, standard errors of parameters, and the variance-covariance matrix for the model currently selected **in the Master Models List. In** addition, this option also provides a legend with a verbal description of the fitted parameters **(see** SURPH Parameter Estimates **).** Estimates are displayed on the scale on which they are estimated by SURPH (e.g., β in the expression $S^{exp(X'\beta)}$), rather than as probabilities.

SURPH Estimates - Displays text window with "integrated probability estimates" for the model currently selected in the **Master Model List. Estimates are displayed** on the probability scale (e.g., $\hat{S}_{ik} = S^{exp(X'_{ik}\beta)}$ **(see** SURPH Probability Estimates **).**

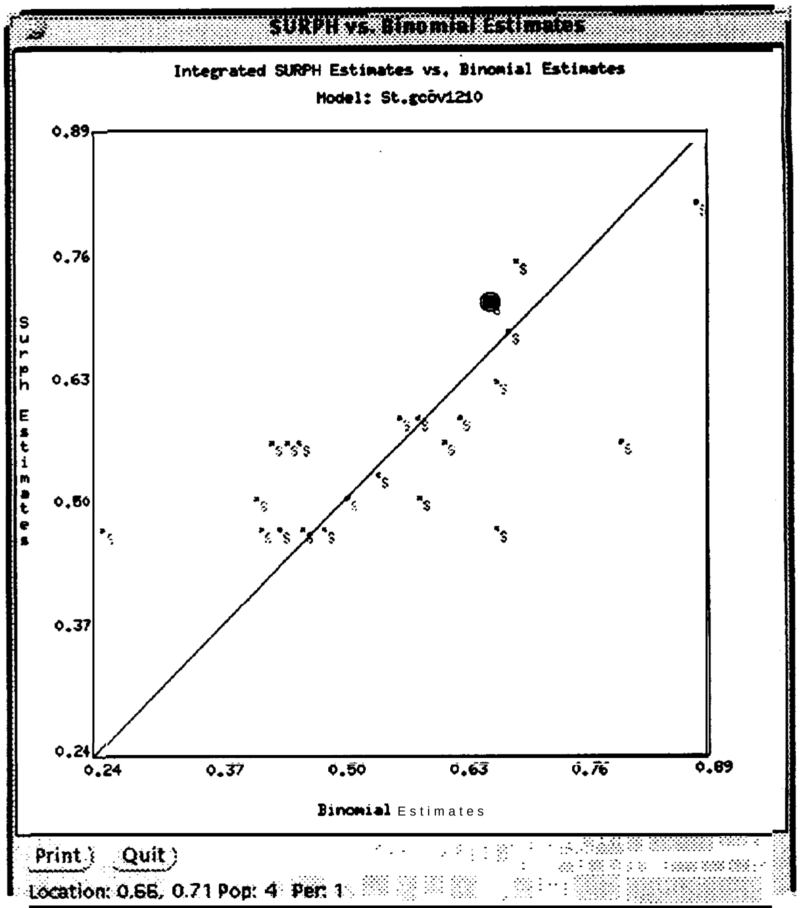
Binomial Estimates - Displays plot of integrated probability estimates of survival probabilities versus the corresponding Binomial Estimates **(see** Models: Binomial Estimates **).**

Survival Curves - Displays graphical representations of group and **individual covariate effects on survival probabilities or on relative risk (see** Survival Graphics **and** Relative Risk Graphics **).**

Residual Plots - Displays graphical representation of model residuals from two perspectives, a) Residuals vs. Capture History, or b) Normal Q-Q Plot **(see** Residual Plot **or** Quantile-Quantile Plot **).**

Change Name - Changes the name of the currently selected model in **the Master Models List (see** Change Model Name

Discard - Discards the currently selected model **in** the Master **Model List.** Models that are discarded are no longer accessible to the user and have to be recomputed if **needed again (see** Confirm Discard Dialog Box **).**



This window **(Models -> Binomial Estimates)** provides a graphical display of the "integrated" **SURPH** probability estimates for **the** model currently selected **in the Master Model List** versus the binomial estimates. Survival ("S's" on the graphic) probabilities are plotted. The control panel along the bottom of the window has buttons **to Print** a copy **of this** window or to **Quit** (dismiss) this window.

This graphic provides a visual diagnostic for assessing the fit of models. Individual' points on the graphic can be selected to identify the period and population they represent. To select a point, move the mouse cursor to that point, then left-click. The point that is displayed above as a small white circle has been selected. The description of the selected point appears in the footer of the window. Here, the selected point is from population 4, period 1. The binomial survival estimate is 0.66, while the SURPH survival estimate is 0.71.

THIS PAGE INTENTIONALLY LEFT BLANK

4.4 SURPH-PC Program

4.4.1 Introduction

Surph-PC is the version of SURPH. 1 that runs under the MS-Windows* operating system. All estimation and hypothesis-testing algorithms are identical to those used in the UNIX version. This portion of Chapter 4 will familiarize the user with cosmetic and procedural differences present within the MS-Windows@ operating system. This chapter does not replace earlier portions of Chapter 4. Instead, this section is intended to **clarify** Chapter 4 withrespect to differences in Surph-PC. The page numbers that follow each new topic indicate where more detailed information may be obtained. The order of the chapter will, where possible, follow the order dictated by the buttons on the **Surph-PC Base Window.** For a full coverage of SURPH.l (i.e., statistical properties, potential uses, examples, and functionality) please read the **SURPH.1** manual.

4.4.2 Differences in Surph-PC

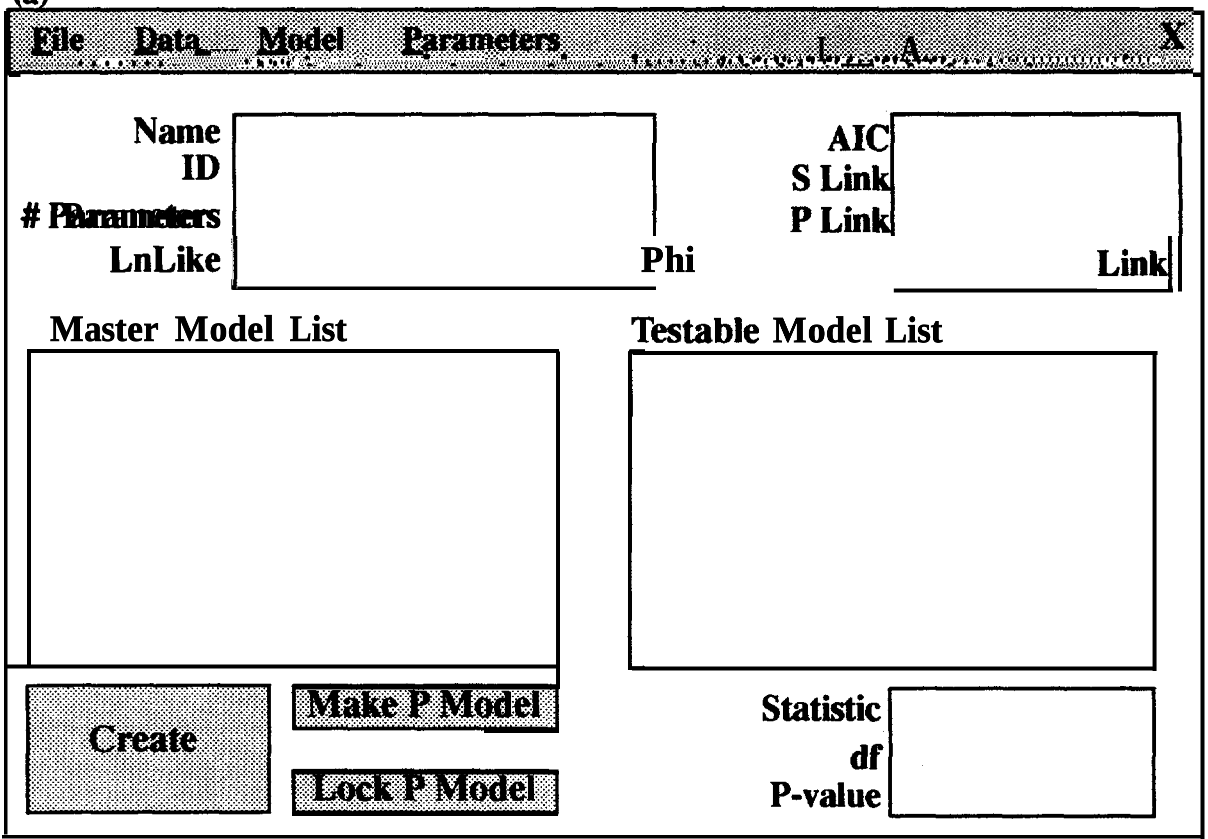
Base Window (pp. 4.27-4.30) **- The Surph-PC Base Window** is cosmetically different from the UNIX version. To begin with, some buttons are **labelled** differently in the two versions. The Surph-PC labels and their UNIX analogues can be found in Table 4.5. Additionally, the upper portion of the **Base Window** looks quite different. In the UNIX version, the **ID,** Name, **LnLik, #** Parameters, S Link, P Link and AIC are displayed for both the model selected in the **Master Model List,** and the model selected **in the Testable Model List. In** Surph-PC, this information is displayed **only** for the model selected **in the Master Model List.**

Likelihood Ratio Tests (pp. 4.21-4.22) **-** In Surph-PC, the results of LRT are displayed in a small box on the lower right-hand portion of the **Surph-PC Base Window.**

Load Data, Load Models, Store Models (pp. 4.33-4.38) **-** As in the UNIX version of SURPH. 1, these functions reside on the **File Pulldown Menu.** However, whereas in the **UNIX** version the user is prompted to type the directory and filename in fields on a pop-up window, in

SURPH Base Window: (a) PC, (b) UNIX

(a)



(b)

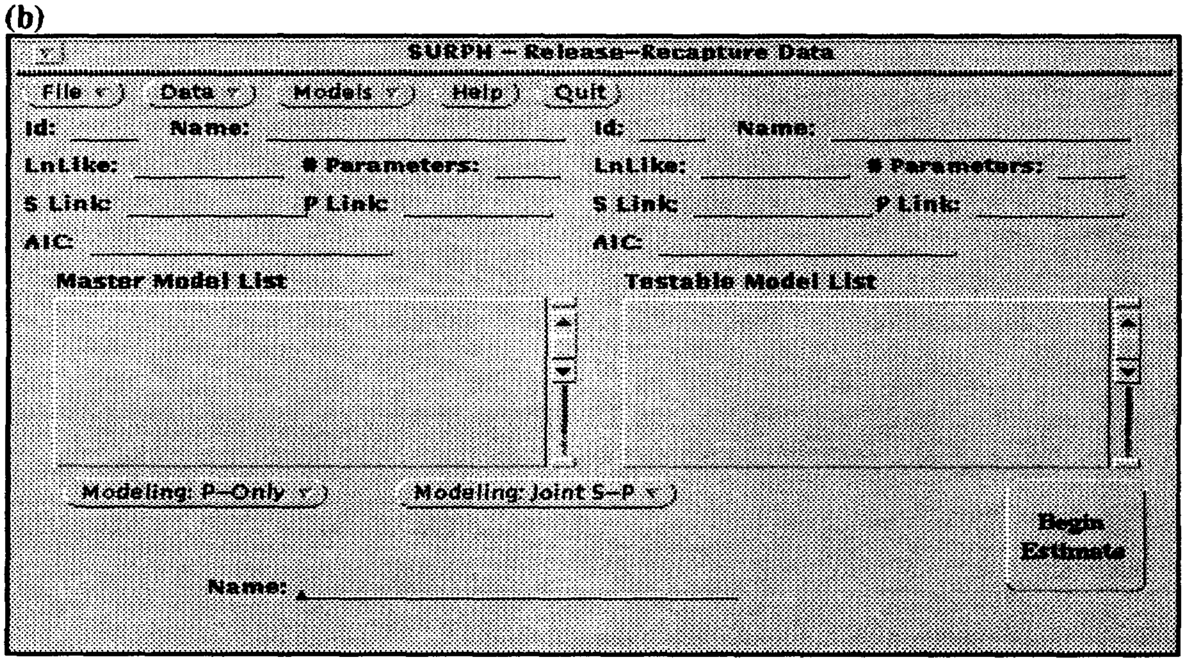


Table 4.5: Base-Window Button Labels. The button label from the UNIX-version of SURPH is given in the left-hand column, along with its PC analogue in the right-hand column.

UNIX	Surph-PC
File	File
Data	Data
Models	Parameters
Help	(Not Available)
Quit	Left-click on "Quit Window" icon **(X)** in upper right corner.
Modeling: P-Only	Parameters
Lock	Lock P Model
Modeling: Joint S-P	Parameters
Begin Estimate	Create

Surph-PC, these operations are preformed through point-and-click mouse operations. During "load" and "store" operations, Surph-PC expects data files to use the **".dat"** suffix, and model files to use the **".mod"** suffix.

ANODEV (pp **4.111 - 4.122) - The ANODEV** option resides on the **File Pulldown Menu in** Surph-PC. **As** with the **Surph-PC Base Window,** some buttons on **the ANODEV Window** have different labels in the Surph-PC version. The Surph-PC labels and their UNIX analogues can be found in Table 4.6.

There are cosmetic changes to **ANODEV as well. In** Surph-PC, the **ANODEV Window has** been split into two separate pop-up windows.The main window contains the information about the models that have been fit, the models that are used to compute **the ANODEV Table,** and shows the computed **ANODEV Table.** The second window is used to create and name additional

Table 4.6: ANODEV Button Labels. The button label from the UNIX-version of SURPH is given in the left-hand column, along with its PC analogue in the right-hand column.

UNIX	**Surph-PC**
Calculate	Compute Table
Print	Print Table
Remove Model	Left-click on Transfer Bar
Transfer Model to Master List	Transfer Model
Quit	Left-click on "Quit Window" icon **(X)** in upper right corner.
Begin Estimation	**Make Model**

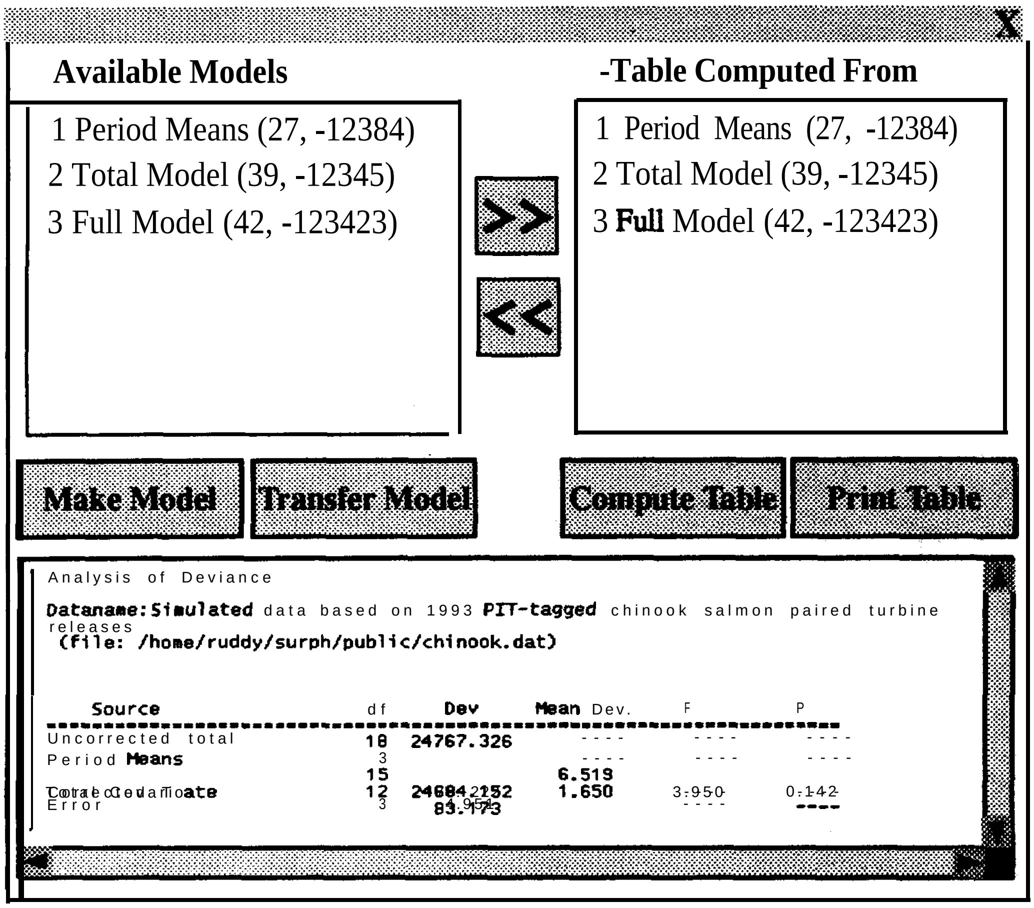
models. In addition, quick-buttons have been added to the second window. **In** the UNIX version of **ANODEV, these** windows are combined.

A procedural difference between Surph-PC and the UNIX version is how models are moved from the **Available Models** region to **the Testable Models** region. In **Surph-PC,** to move a model from one region to the other, select the model (left-click), then left-click on the transfer bars (i.e., ).

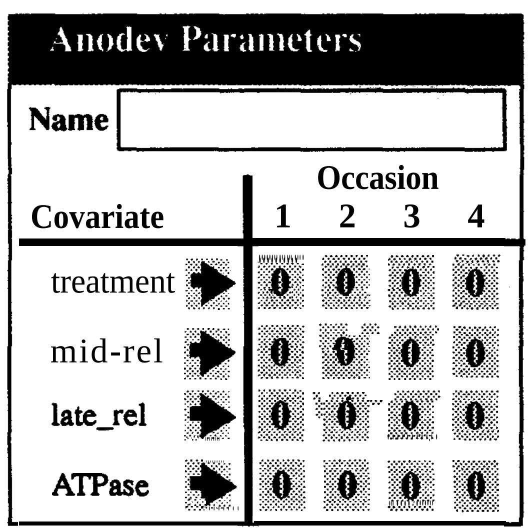
Individual Covariates (pp. 4.47-4.56) **-** As in the UNIX version of SURPH.l, this function resides on the **Data Pulldown Menu.** A histogram of the data or a cumulative distribution plot of the data may be displayed using this selection.

The windows that are displayed for the PC version have some minor cosmetic differences compared to their UNIX analogues. First, the positioning of the various components displayed has been changed. However, all options that are available in the UNIX version are available in Surph-PC. Second, the slide bar that was used to change the "Occasion" designation in the UNIX version has been replaced by a **pulldown** menu **in** Surph-PC. Finally, the **Quit Button has** been

(a)


Available Models
1 Period Means (27, -12384)
2 Total Model (39, -12345)
3 Full Model (42, -123423)
-Table Computed From
1 Period Means (27, -12384)
2 Total Model (39, -12345)
3 Full Model (42, -123423)
Make Model
Transfer Model
Compute Table
Print Table
Analysis of Deviance
Dataname: Simulated data based on 1993 PIT-tagged chinook salmon paired turbine releases
(file: /home/ruddy/surph/public/chinook.dat)
Source
df
Dev
Mean Dev.
F
P
Uncorrected total
18
24767.326
Period Means
3
15
6.519
12
1.650
3.950
0.142
Error
3
83.173

(b)


Anodev Parameters
Name
Occasion
Covariate
1
2
3
4
treatment
0
0
0
0
mid-rel
0
0
0
0
late_rel
0
0
0
0
ATPase
0
0
0
0

Analysis of Deviance

Calculate | Print | Remove Model | Transfer Model to Master List | Quit

Available Models

1 Period Means (27, -12384)
2 Total Model (39, -12345)
3 Full Model (42, -12342)

Table Computed From

1 Period Means (27, -12384)
2 Total Model (39, -12345)
3 Full Model (42, -12342)

Effect:	Period 1	Period 2	Period 3
treatment	Off	Off	Off
mid_release	Off	Off	Off
late_release	Off	Off	Off
ATPase	Off	Off	Off

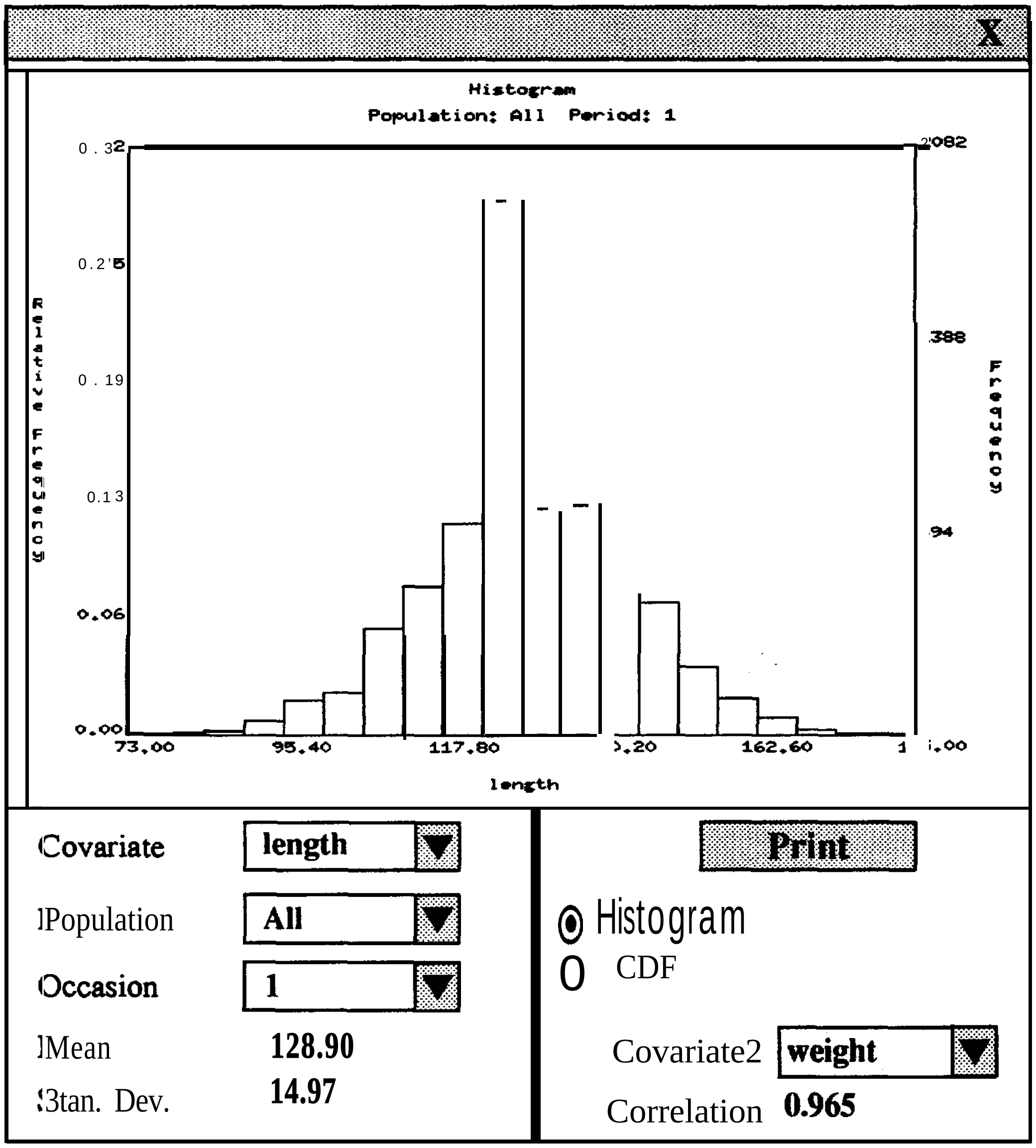
Name:

Begin Estimation

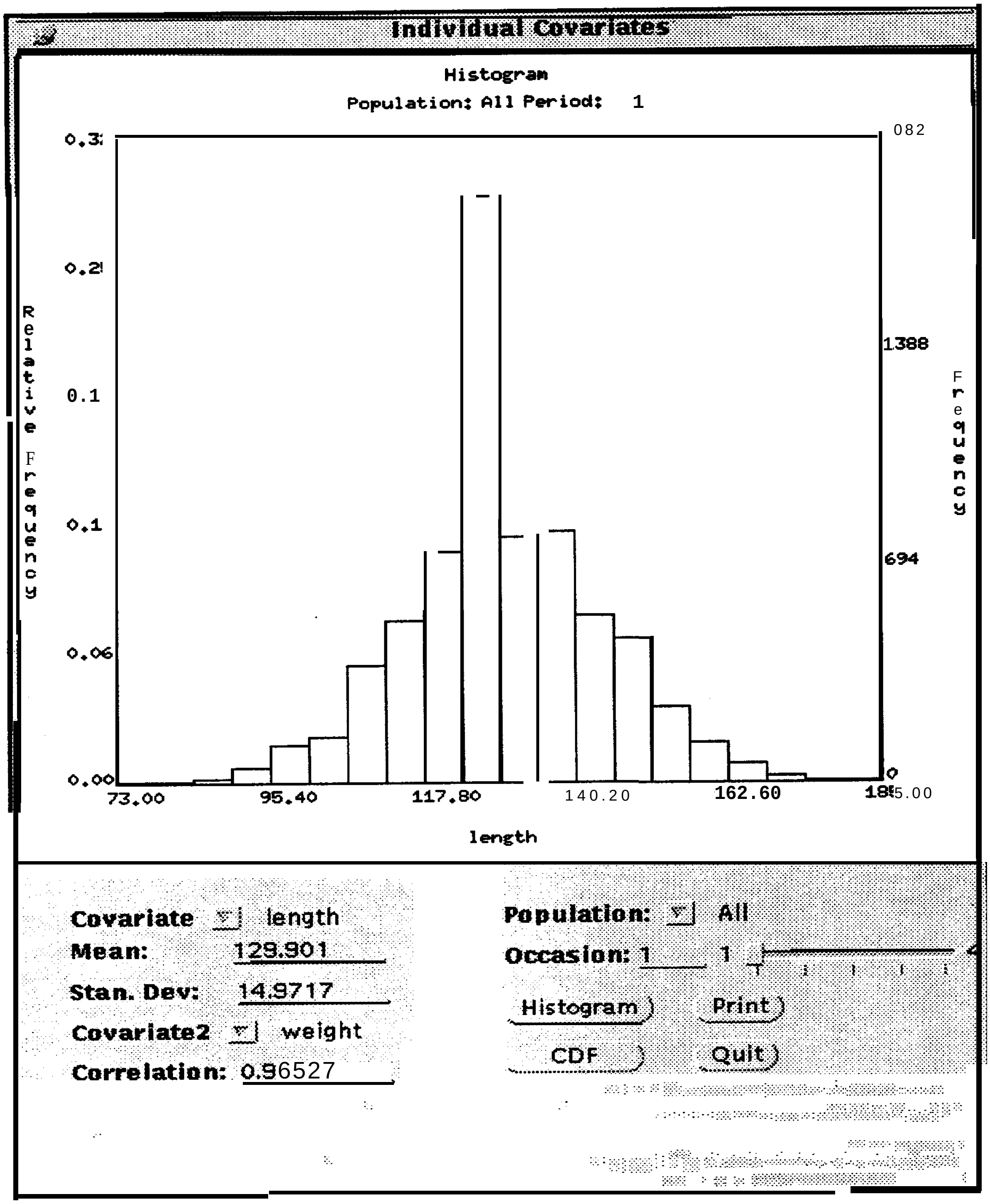
Analysis of Deviance

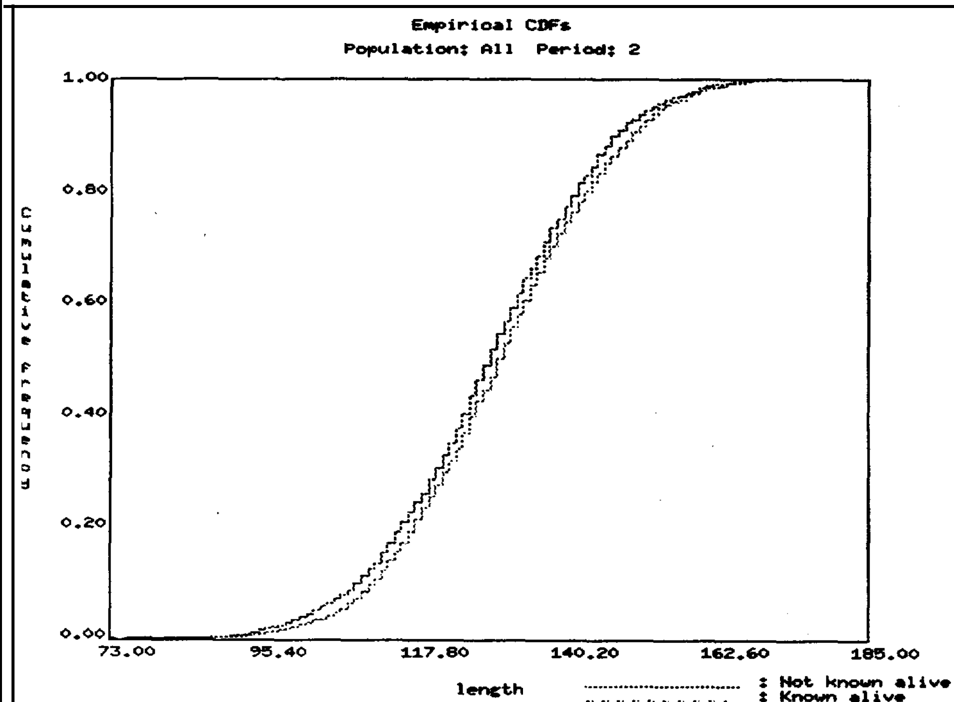
Dataname: Simulated data based on 1993 PIT-tagged chinook salmon palred turbine releases
(file: /home/ruddy/surph/public/chinook.dat)

Source	df	Dev	Mean Dev.	F	P
Uncorrected total	18	24767.326	----	----	----
Period Means	3	93.173	----	----	----
Corrected Total	15	24684.152	-we-	---e	----
Total Covariate	12	78.222	6.519	3.950	0.142
Error	3	4.951	1.650	----	----


Histogram
Population: All Period: 1
Relative Frequency
Frequency
0.32
0.25
0.19
0.13
0.06
0.00
73.00
95.40
117.80
162.60
length
Covariate
length
Population
All
Occasion
1
Mean
128.90
Stan. Dev.
14.97
Print
Histogram
CDF
Covariate2
weight
Correlation
0.965

SURPH Individual Covariates: Histogram





Covariate	length	**Print**	
Population	all	Histogram	
Occasion		CDF	
Mean	128.90	Covariate2	weight
Stan. Dev.	14.97	Correlation	0.965

Kolmogorov-Smirnov Test X

Sample Sizes

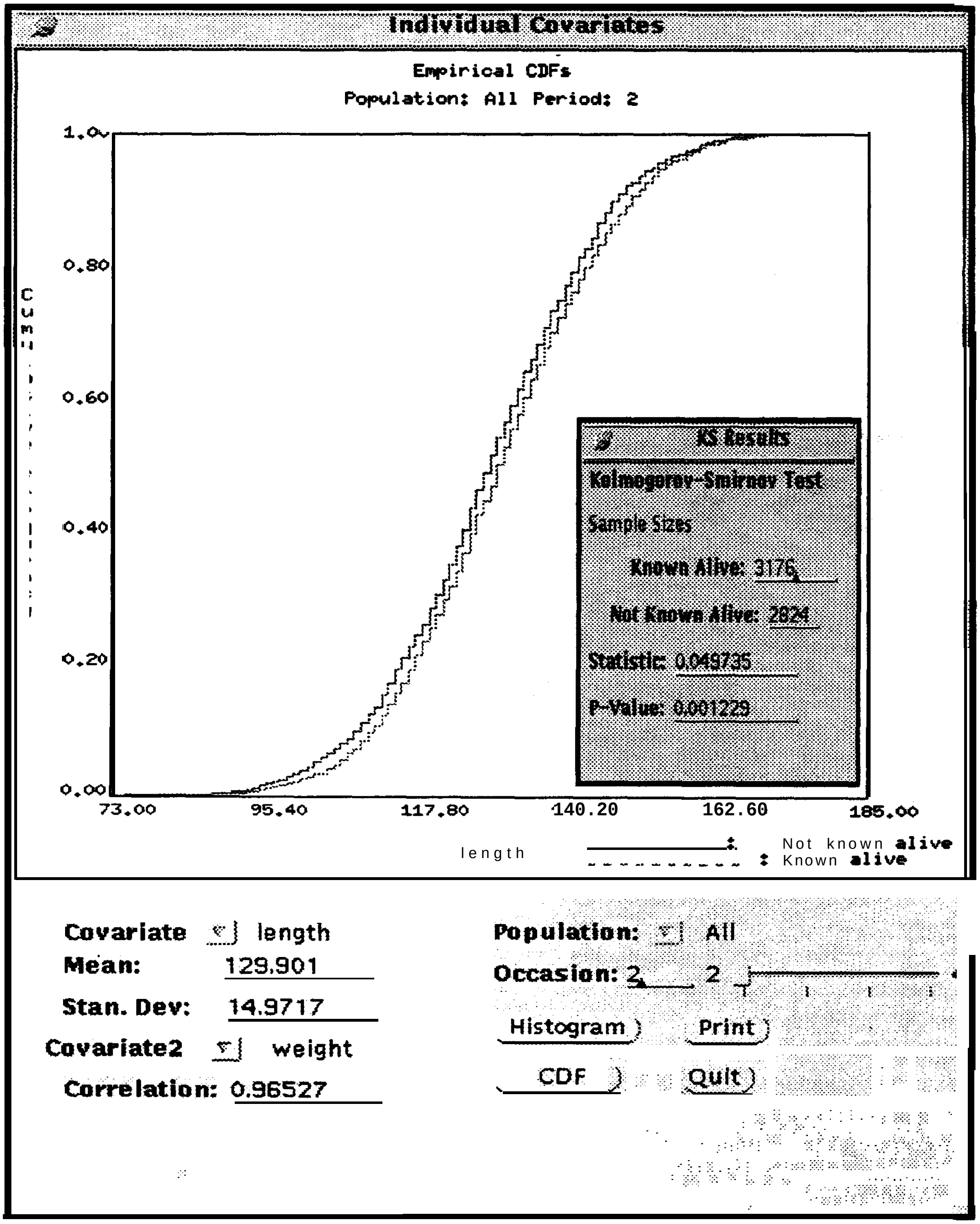
Known Alive 3176

Not Known Alive 2824

statistic 0.0497

P-Value 0.001

SURPH Individual Covariates: CDF



removed. In Surph-PC the user must left-click on the “Quit Window” icon (i.e., X) in the upper right-hand corner to quit.

Capture Modeling (pp. **4.59-4.90) -** One obvious change **in** Surph-PC **is** that **all** of the **Button** Pads are accessed using **the Parameters Pulldown Menu. Thus, in** Surph-PC, the difference between P-Only Modeling and Joint S-P Modeling is defined by which button is depressed prior to estimation of the model (i.e., **Create vs. Make P Model),** not which Button Pads are accessed. To initiate P-Only Modeling in Surph-PC, the user accesses the Capture Button Pads from the **Parameters Pulldown Menu.** The user then selects the **Make P Model** button on the **Surph-PC Base Window** to estimate the model. **In UNIX,** to initiate **P-Only** modeling, the user would select the Capture Button Pads from the **Modeling: P-Only Pulldown Menu.**

Renaming Models (pp. 4.173-4.174) **- In** the UNIX version, a Name field is provided on the **SURPH Base Window** prior to estimation. **In** Surph-PC, the name is initially fixed. The name can only be changed after estimation using the **Rename** command **(Model -> Rename).**

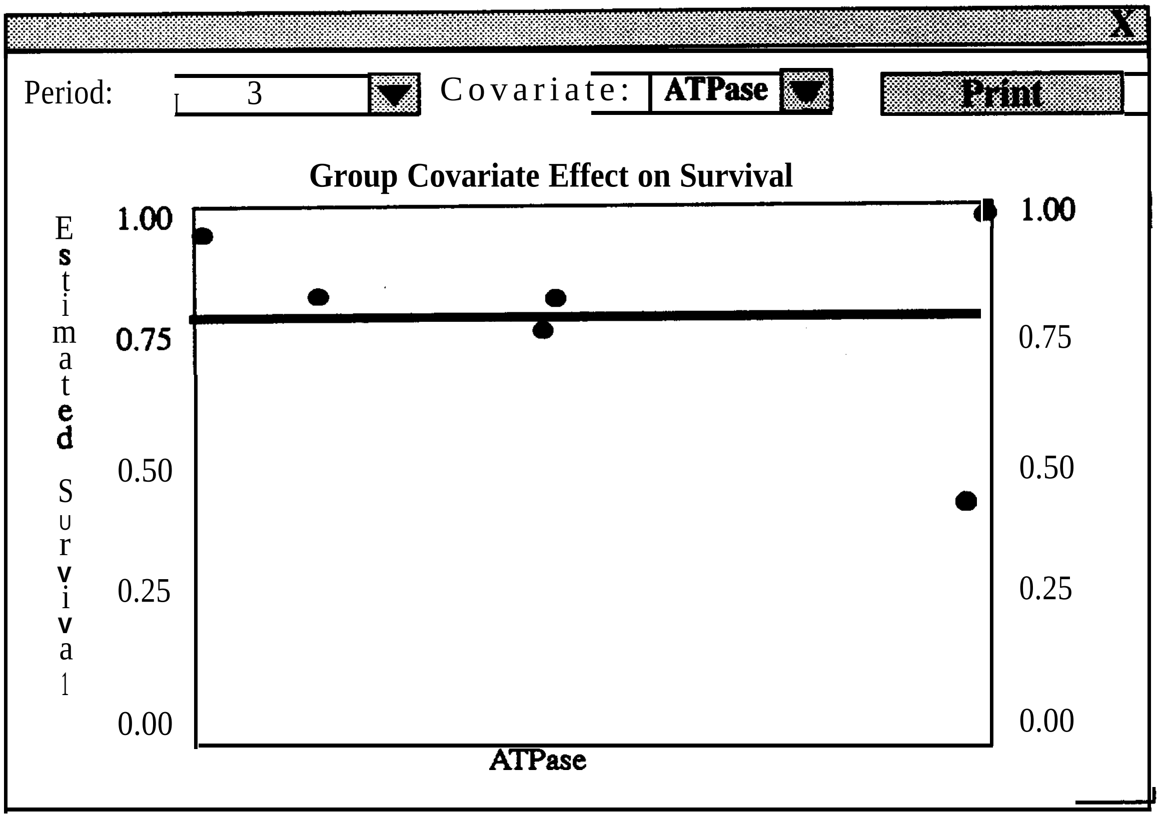
Survival, Capture and Product Curves (pp. 4.139-4.168) **-** As in the **UNIX** version of SURPH. 1, these functions reside on the **Model Pulldown Menu. The** graphics that are displayed (i.e., group and individual curves) have some minor cosmetic differences. Only the initial pop-up window is displayed for each covariate type (i.e., group or individual) as all the cosmetic changes have occurred to the pop-up window itself, not the graphic within the pop-up window.

For the window that displays the group-covariate curves, the buttons have been moved. **In** Surph-PC, the buttons are located on the upper portion of the pop-up window, whereas in the Unix version of **SURPH,** the buttons are located on the lower portion of the pop-up window. Otherwise the pop-up window is almost identical.

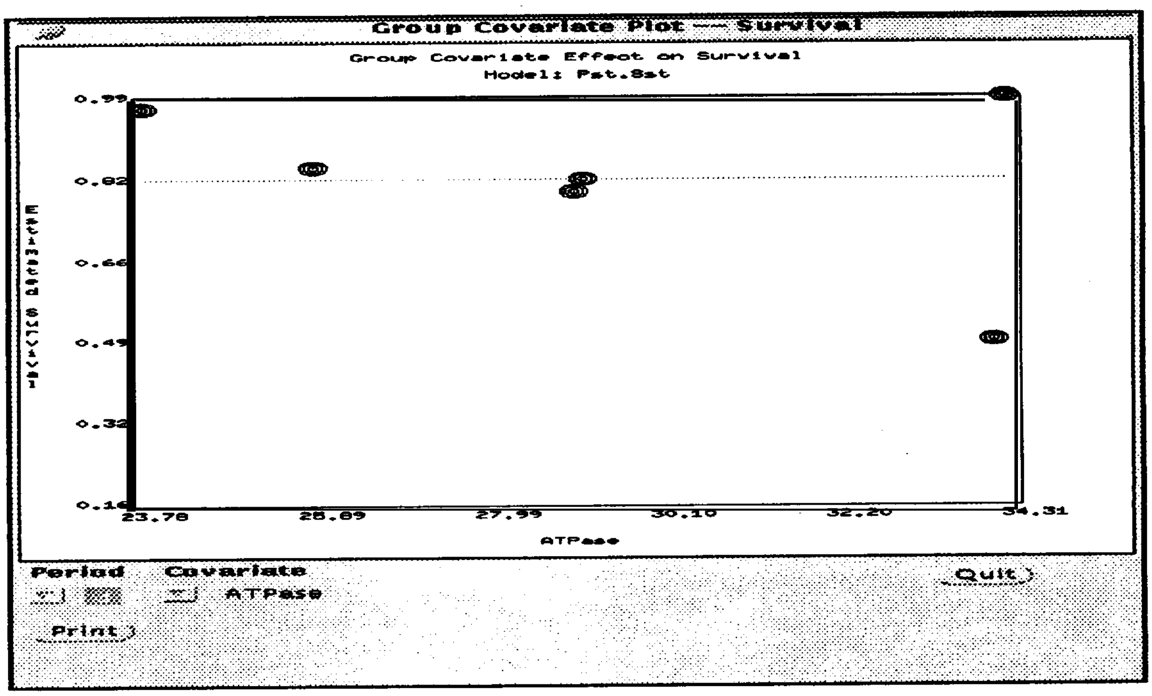
For the window that displays the individual-covariate curves, there are only cosmetic differences. Primarily, the location of the buttons that select between the various SURPH-curve options has been altered. **In** Surph-PC, the buttons are located on the left-hand side of the window,

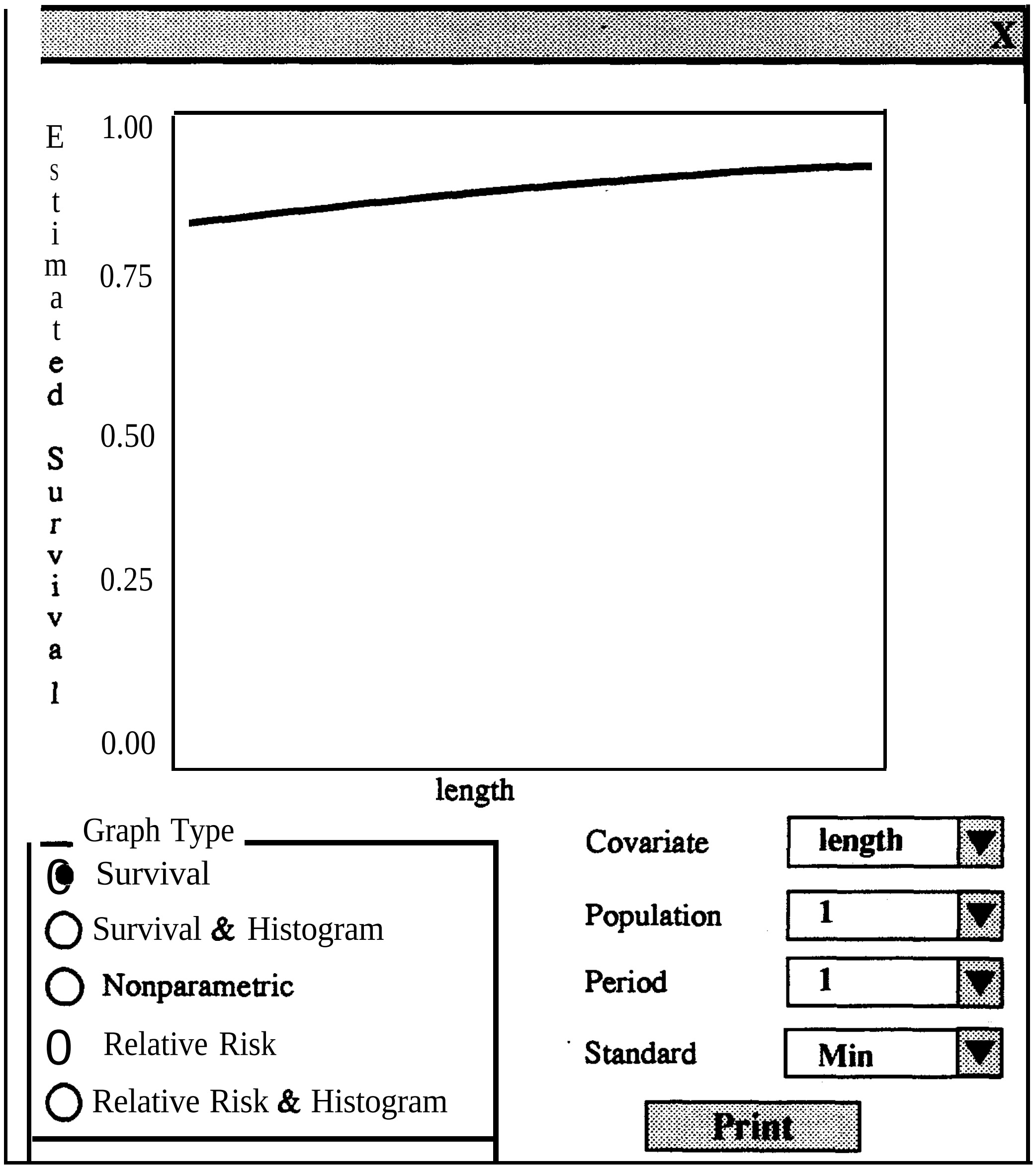
Group Covariates Curves: (a) PC, (b) UNIX

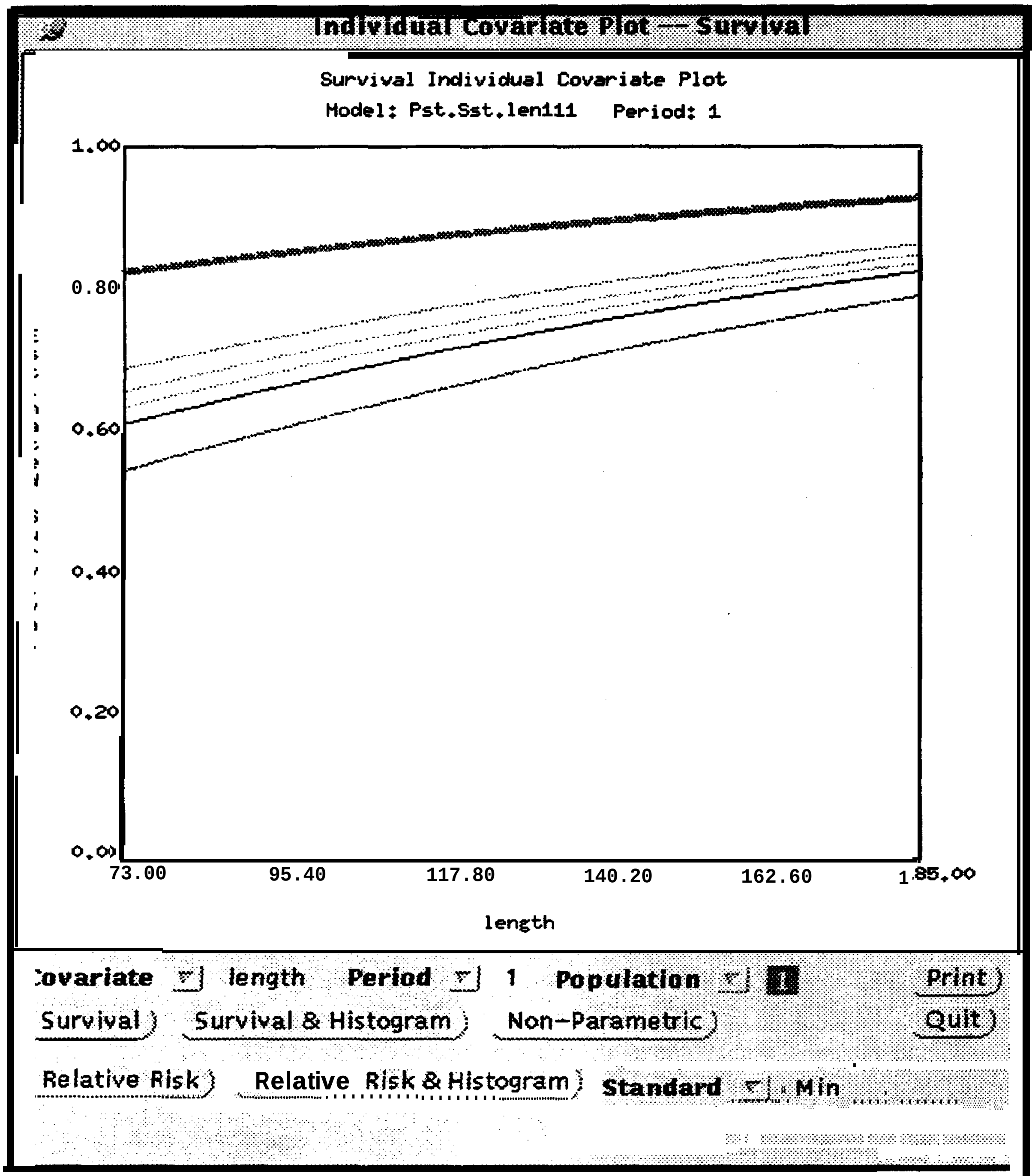
(a)



(b)




X
Estimated Survival
1.00
0.75
0.50
0.25
0.00
length
Graph Type
Survival
Survival & Histogram
Nonparametric
Relative Risk
Relative Risk & Histogram
Covariate
length
Population
1
Period
1
Standard
Min
Print


Individual Covariate Plot — Survival
Survival Individual Covariate Plot
Model: Pst.Sst.len111 Period: 1
1.00
0.80
0.60
0.40
0.20
0.00
73.00
95.40
117.80
140.20
162.60
185.00
length
Covariate
length
Period
1
Population
Print
Survival
Survival & Histogram
Non-Parametric
Quit
Relative Risk
Relative Risk & Histogram
Standard
Min

and label is adjacent to the button. On the UNIX version, the label of the graphic to be displayed is atop the button. **All** of **the** buttons except for the **Quit Button** are present. Similarly, the **pulldown** menus have been relocated. In Surph-PC, all pulldowns are located on the right-hand side of the window, whereas in the UNIX version, the **pulldown** menus were located above the graphics buttons.

Literature Cited

AKAIKE, H. 1973. Information theory and an extension of the maximum likelihood principle. *In* B.N. Prtrov and F. Csaki [ed.] 2nd International Symposium on Information Theory. Akademiai Kiado, Budapest.

BANTOK, C. R., and J. A. BAYLEY. 1973. Visual selection for shell size in *Cepaea* (Held.). J. Anim. Ecol. 42: 247-261.

BROWNIE, C., D. R. ANDERSON, K. P. BURNHAM, and D. S. ROBSON. 1985. Statistical inference from band recovery data--A handbook. U.S. Fish and Wildlife Service Resource Publication No. 156,305 p.

BURNHAM, K. P., D. R. ANDERSON, G. C. WHITE, C. BROWNIE, and K. H. POLLOCK. 1987. Design and analysis methods for fish survival experiments based on release-recapture. Am. Fish. Soc. Monograph No. 5,437 p.

BYRON, E. R. 1982. The adaptive significance of calanoid copepod pigmentation: a comparative and experimental analysis. Ecology. 63: 187 1- 1886.

CLOBERT, J., and J. D. LEBRETON. 1985. Dépendance de facteurs de milieu dans les estimations de taux de survie par capture-recapture. Biometrics. 41: 1031-1037.

CLOBERT, J., J. D. LEBRETON, and D. ALLAINE. 1987. A general approach to survival rate estimation by recaptures or resightings of marked birds. Ardea. 75: 133-142.

CLOBERT, J., J. D. LEBRETON, M. CLOBERT-GILLET, and H. COQUILLART. 1985. The estimation of survival in bird populations by recaptures or sightings of marked individuals. *In* B. J. T. Morgan and P. M. North [ed.] Lecture Notes in Statistics 29. Statistics in Ornithology. Berlin: Springer-Verlag, pp. 197-213.

CLOBERT, J., C. M. PERRINS, R. H. MCCLEERY, and A. G. GOSLER. 1988. Survival rate in the Great Tit *Parus major* in relation to sex, age, and immigration status. J. Anim. Ecol. 57: 287-306.

CORMACK, R. M. 1964. Estimates of survival from the sighting of marked animals. Biometrika. 5 1: 429-438.

CORMACK, R. M. 1970. Statistical appendix. *In* R. A. Fordham Mortality and population change of Dominican gulls in Wellington, New Zealand. J. Anim. Ecol. 39: 13-27.

COX, D. R. 1958. Planning of Experiments. New York: Wiley, 308 p.

COX, D. R. and E. J. SNELL. 1989. The Analysis of Binary Data. 2nd Edition. Chapman and Hall, London. 236 p.

DICESNOLA, A. P. 1904. *Preliminary* notes on *the* protective value of colour *in Mantis religiosa.* Biometrika. 3: 58-59.

ENDLER, J. A. 1986. Natural Selection in the Wild. Princeton University Press, Princeton, NJ. 337 p.

FEDERER, W. T. 1973. Statistics and Society. New York: Marcel Dekker, 399 p.

FEIGL, P., and M. ZELEN. 1965, Estimation of exponential survival probabilities with concomitant information. Biometrika. 2 1: 826-38.

FISHER, R. A. 1922. On the mathematical foundations of theoretical statistics. Philos. Trans. R. Soc. London A, 222:309-368.

FISHER, R. A. 1925. Theory of statistical estimation. Proc. Camb. Phil. Soc. 22:700-725.

FISHER, R. A. 1947. The Design of Experiments, 4th Edition. New York: Hafner, 240 p.

FLETCHER, R. 1970. A new approach to variable metric algorithms. Computer Journal, 13:317-322.

FORDHAM, R. A., and R. M. CORMACK. 1970. Mortality and population change of dominican gulls in Wellington, New Zealand, with a statistical appendix. J. Anim. Ecol. 39: 13-27.

HOEL, P. G., S. C. PORT, C. J. STONE. 197 1. Introduction to Probability Theory. Houghton Mifflin, Boston, MA. 258 p.

HOFFMANN, A. 1993. Quantifying selection in wild populations using known-fate and mark-recapture designs. Ph.D. dissertation, University of Washington, Seattle, WA. 259 p.

HOGG, R. V., and E. A. TANIS. 1983. Probability and Statistical Inference, Second Edition. Macmillan Publishing Co., Inc., New York, NY. 533 p.

HOSMER, D. W. and S. LEMESHOW, JR. 1989. Applied Logistic Regression. New York: Wiley, 307 p.

JOLLY, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration--stochastic model. Biometrika. 52: 225-47.

KALBFLEISCH, J. D., and R. L. PRENTICE. 1980. The statistical analysis of failure time data. John Wiley & Sons, New York, NY. 321 p.

KAPLAN, E. L., and P. MEIER. 1958. Nonparametric estimation from incomplete observations. J. Amer. Stat. Assoc. 53: 457-481.

LACK, D. 195 1. Population ecology in birds. Proc. Xth Int. Om. Congr. Uppsala. pp. 409-48.

LAWLESS, J.F. 1982. Statistical model and methods for lifetime data. New York: John Wiley, 580 p.

LEBRETON, J. D., K. P. BURNHAM, J. CLOBERT, and D. R. ANDERSON. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecol. Monogr. 62(1): 67-118.

MCCULLAGH, P., and J. A. NELDER. 1983. Generalized Linear Models. 1st edition. John Wiley & Sons, New York, NY. 261 p.

NORTH. P. M., and B. J. T. MORGAN. 1979. Modeling heron survival using weather data. Biometrics. 35: 667-68 1.

POLLOCK, K. H., J. E. I, and J. D. NICHOLS. 1984. The use of auxiliary variables in capture-recapture and removal experiments. Biometrics 40:329-340.

POLLOCK, K. H., S. R. WINTERSTEIN, C. M. BUNCK, and P. D. CURTIS. 1989a. Survival analysis in telemetry studies: The staggered entry design. J. Wildl. Manage. 53: 7-15.

POLLOCK, K. H., S. R. WINTERSTEIN, and M. J. CONROY. 1989b. Estimation and analysis of survival distributions for radio-tagged animals. Biometrics. 45: 99-109.

PRESS, WH, B. P. FLANNERY, S. A. TEUKOLSKY, AND W. T. VETTERLING. 1986. Numerical Recipes. Cambridge: Cambridge University Press, 818 p.

SAKAMOTO, Y., M. ISHIGURO, and G. KITAGAWA. 1987. Akaike information criterion statistics. KTK Scientific Publishers, Tokyo. 290 p. (Transl. from Japanese by Johoryo Tokeigaku.)

SANDLAND, R. L., and P. KIRKWOOD. 1981. Estimation of survival in marked populations with possibly dependent sighting probabilities. Biometrika. 68: 53 1-541.

SCHULZE, P. C., and C. L. FOLT. 1990. Food resources, survivorship, and reproduction of the omnivorous calanoid copepod *Epischura lacustris*. Ecology. 7 l(6): 2224-2240.

SCHWARZ, C. J. 1988. Post-release stratification and migration models in band-recovery and capture-recapture models. Ph.D. dissertation, University of Manitoba, Canada, 384 p.

SEBER, G. A. F. 1962. The multi-sample single recapture census. Biometrika. 49: 330-49.

SEBER, G. A. F. 1965. A note on the multiple recapture census. Biometrika. 52: 249-52.

SEBER, G. A. F. 1970. Estimating time-specific survival and reporting rates for adult birds from band returns. Biomettika. 57: 313-18.

SEBER, G. A. F. 1972. Estimating survival rates from bird-band returns. J. Wild& Manage. 36: 405-13.

SEBER, G. A. F. 1982. The Estimation of Animal Abundance and Related Parameters. Griffin, London. 654 p.

SKALSKI, J. R., A. HOFFMANN, and S.G. SMITE. 1993. Testing the significance of individual- and cohort-level covariates in animal survival studies, p. 9-28. *In* J. D. Lebreton & P. M. North [ed.] Marked Individuals in the Study of Bird Population. Birkhauser Verlag, Basel.

SKALSKI, J. R., and D. S. ROBSON. 1992. Techniques for wildlife investigations: Design and analysis of capture data. Academic Press, San Diego, CA. 237 p.

SMITH, S. G. 1991. Assessing hazards in wild populations using auxiliary variables in&g-release models. Ph.D. dissertation, University of Washington, Seattle, WA. 351 p.

SMITH, S. G., and J. R. SKALSKI. Assessing the statistical significance of group covariate effects on survival using analysis of deviance of release-recapture data. J. Applied. Ecol. (accepted).

VAESSEN, W. 1984. UBC NLI? nonlinear function optimization. Computing Centre, Univ. British Columbia.

WALTON, I. 1653. The compleat angler. Repr. Everyman's Lib. London, 1953,215 p.

WELTY, J. C. 1982. The Life of Birds. New York: Saunders College Publisher, 754 p.

WHITE, G. C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. J. Wildl. Manage. 47: 716-728.

WHITE, G. C. and R. A. **GARROTT**. 1990. Analysis of Wildlife Radio-Tracking Data. Academic Press, San Diego, CA. 383 p.